

**SEARCH REQUEST FORM**

Scientific and Technical Information Center

Requester's Full Name: \_\_\_\_\_ Examiner #: \_\_\_\_\_ Date: \_\_\_\_\_

Art Unit: \_\_\_\_\_ Phone Number 30 \_\_\_\_\_ Serial Number: \_\_\_\_\_

Mail Box and Bldg/Room Location: \_\_\_\_\_ Results Format Preferred (circle): PAPER DISK E-MAIL

If more than one search is submitted, please prioritize searches in order of need.

Please provide a detailed statement of the search topic, and describe as specifically as possible the subject matter to be searched. Include the elected species or structures, keywords, synonyms, acronyms, and registry numbers, and combine with the concept or utility of the invention. Define any terms that may have a special meaning. Give examples or relevant citations, authors, etc. if known. Please attach a copy of the cover sheet, pertinent claims, and abstract.

Title of Invention: \_\_\_\_\_

Inventors (please provide full names): \_\_\_\_\_

Earliest Priority Filing Date: \_\_\_\_\_

\*For Sequence Searches Only: Please include all pertinent information (parent, child, divisional, or issued patent numbers) along with the appropriate serial number.

**STAFF USE ONLY**

Searcher: D. Schreiber  
 Searcher Phone #: 308-4292  
 Searcher Location: CMI 12C14  
 Date Searcher Prepr'd: \_\_\_\_\_  
 Date Implemented: 12/10  
 Searcher Prep & Review Time: 25  
 Clerical Prep Time: \_\_\_\_\_  
 Other Time: ?

Type of Search	Vendors and cost where applicable
NA Sequence (#): <u>7</u>	STN: _____
AA Sequence (#): <u>1</u>	Dialog: _____
Structure (#): _____	Questel Orbit: _____
Bibliographic: _____	Orbit: _____
Citation: _____	Lexis/Nexis: _____
Fulltext: _____	Sequence Systems: <u>Cambridge</u>
Patent Family: _____	WWW: Internet: _____
Other: _____	Other: _____

**THIS PAGE BLANK (USPTO)**

GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

## OM protein - protein search, using sw model

Run on: November 5, 2001, 12:56:58 ; Search time 23.73 Seconds

(without alignments)  
18.729 Million cell updates/sec

**Title:** US-08-894-356C-21  
**Perfect score:** 40 DFGWIK 6  
**Sequence:** BLOSUM62  
**Scoring table:** Gapext 10.0 , Gapext 0.5  
**Searched:** 522463 seqs, 74073290 residues  
**Total number of hits satisfying chosen parameters:** 522463  
**Minimum DB seq length:** 0  
**Maximum DB seq length:** 200000000  
**Post-processing:** Minimum Match 0%, Maximum Match 100%, Listing first 45 summaries

## Database :

A\_Geneseq\_1101:\*

- 1: /SIDS8/gcadata/geneseq/geneseq/geneseqp/A11980.DAT:\*
- 2: /SIDS8/gcadata/geneseq/geneseqp/A11981.DAT:\*
- 3: /SIDS8/gcadata/geneseq/geneseqp/A11982.DAT:\*
- 4: /SIDS8/gcadata/geneseq/geneseqp/A11983.DAT:\*
- 5: /SIDS8/gcadata/geneseq/geneseqp/A11984.DAT:\*
- 6: /SIDS8/gcadata/geneseq/geneseqp/A11985.DAT:\*
- 7: /SIDS8/gcadata/geneseq/geneseqp/A11986.DAT:\*
- 8: /SIDS8/gcadata/geneseq/geneseqp/A11987.DAT:\*
- 9: /SIDS8/gcadata/geneseq/geneseqp/A11988.DAT:\*
- 10: /SIDS8/gcadata/geneseq/geneseqp/A11989.DAT:\*
- 11: /SIDS8/gcadata/geneseq/geneseqp/A11990.DAT:\*
- 12: /SIDS8/gcadata/geneseq/geneseqp/A11991.DAT:\*
- 13: /SIDS8/gcadata/geneseq/geneseqp/A11992.DAT:\*
- 14: /SIDS8/gcadata/geneseq/geneseqp/A11993.DAT:\*
- 15: /SIDS8/gcadata/geneseq/geneseqp/A11994.DAT:\*
- 16: /SIDS8/gcadata/geneseq/geneseqp/A11995.DAT:\*
- 17: /SIDS8/gcadata/geneseq/geneseqp/A11996.DAT:\*
- 18: /SIDS8/gcadata/geneseq/geneseqp/A11997.DAT:\*
- 19: /SIDS8/gcadata/geneseq/geneseqp/A11998.DAT:\*
- 20: /SIDS8/gcadata/geneseq/geneseqp/A12000.DAT:\*
- 21: /SIDS8/gcadata/geneseq/geneseqp/A12001.DAT:\*
- 22: /SIDS8/gcadata/geneseq/geneseqp/A12001.DAT:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result	Score	Query Length	DB ID	Description
1	40	100.0	6 17 AAW04728	Aromatic acyl tran
2	40	100.0	7 22 AAE00248	Peptide fragment #
3	40	100.0	287 21 AAG07179	Arabidopsis thalia
4	40	100.0	287 21 AAG42296	Arabidopsis thalia
5	40	100.0	293 21 AAG0708	Arabidopsis thalia
6	40	100.0	293 21 AAG42295	Arabidopsis thalia
7	40	100.0	302 22 AAE00230	Arabidopsis thalia
8	40	100.0	302 22 AAE00231	Arabidopsis thalia
9	40	100.0	302 22 AAE00234	Arabidopsis thalia
10	40	100.0	302 22 AAE00235	Arabidopsis thalia
11	40	100.0	302 22 AAE00238	Arabidopsis thalia

## ALIGNMENTS

RESULT	1	AAW04728	AAW04728 standard; Protein; 6 AA.
ID	AAW04728	XX	AAW04728;
DT	06-FEB-1997	( first entry)	
DE			Aromatic acyl transferase peptide fragment.
XX			Aromatic acyl transferase; transformation; anthocyanin pigment; plants; acylation; colour; tone; colouration; colour change; Gentiana triflora; Petunia hybrida; Perilla ocimoides; Scenecio cruentus; Lavandula angustifolia.
OS			Gentiana triflora var. japonica (Clone pgATA4).
PN			N09525500-A1.
XX			
PD	22-AUG-1996.		
XX			
EF	15-FEB-1996;	96WO-JP00348.	
XX			
FR	30-JAN-1996;	96JP-0046534.	
FR	17-FEB-1995;	95JP-0061159.	
FR	29-JUN-1995;	95JP-0196915.	
XX			
PA	(SUNR ) SUNTORY LTD.		
XX			
PI	Asahikari T., Fujiwara H., Fukui Y., Kusumi T., Mizutani M;		
PI	Nakao M., Tanaka Y., Yonekura K;		
XX			
DR	W21; 1996-393401/39.		
XX			

PT DNA coding for aromatic acyl transferase - for transforming plants which produce anthocyanin pigments and thus altering colour tone,  
e.g. of flowers  
XX  
PS Claim 2; Page 81; 94pp; Japanese.  
XX  
CC vectors containing DNA fragments encoding proteins of plant origin with aromatic acyl transferase activity may be used to transform plants which produce anthocyanin pigments. The aromatic acyl transferase acylates the pigments in the flower resulting in colour tone changes and allowing new colourations to be produced. Six specific DNA sequences encoding aromatic acyl transferase from different plants are described in AAT77308-137313. This peptide fragment was isolated from three of the six clones and was used to synthesize a degenerate primer (AAT37314) which was then used to identify other aromatic acyl transferase encoding clones.  
XX  
SQ Sequence 6 AA;

Query Match 100.0%; Score 40; DB 17; Length 6;  
Best Local Similarity 100.0%; Pred. No. 4.3e+05; Mismatches 0;  
Matches 6; Conservative 0; Indels 0; Gaps 0;  
Qy 1 DFGNGK 6  
ID 1 |||||  
AC 1 dfgngk 6  
DB 1 dfgngk 6

## RESULT 2

AAE00248

DE AAE00248 standard; peptide; 7 AA.

XX

AC AAE00248;

XX

DT 13-JUN-2001 (first entry)

XX

DE Peptide fragment #3 to construct degenerate reverse PCR primer AT-REVI.

XX

KW Transacylase; taxol; paclitaxel biosynthesis; taxoid;

XX

OS Unidentified.

XX

PN WO20012356-A2.

XX

PD 05-APR-2001.

XX

PP 29-SEP-2000; 2000WO-US277006.

XX

PR 30-SEP-1999; 99US-0411145.

XX

PR 07-DEC-1999; 99US-0457046.

XX

PA (UNIV ) UNIV WASHINGTON STATE RES FOUND.

XX

PI Croteau RB, Walker KD, Schoendorf A, Wildung MR;

XX

DR WPI; 2001-245004/25.

XX

PT New transacylase enzymes, useful for the high yield production of paclitaxel, related taxoids and useful intermediates in the paclitaxel biosynthetic pathway

PT

PS Disclosure; Page 40; 162pp; English.

XX

CC The present sequence is a peptide fragment used for constructing degenerate reverse PCR primer AT-REVI. The primers are useful for generating probes which are are useful for the identification of (nucleic acid sequences encoding) transacylases. The probes isolated from the Taxus genus are useful for the synthetic production of taxol(TM) and related taxoids, as well as intermediates in the paclitaxel biosynthetic pathway. They can also be used for the creation of transgenic organisms that either produce the transacylases for

CC subsequent in vitro use, or produce the transacylases in vivo. The (nucleic acids encoding) transacylases are also useful for creating specific binding agents that recognise the corresponding transacylases. Binding agents include (fragments of) antibodies or any other agent capable of specifically binding to the groups on the proteins.

## RESULT 3

AAG07709

ID AAG07709 standard; Protein; 287 AA.

XX

AC AAG07709;

XX

DT 17-OCT-2000 (first entry)

XX

DE Arabidopsis thaliana protein fragment SEQ ID NO: 4964.

XX

KW Protein identification; signal transduction pathway; metabolic pathway; hybridisation assay; genetic mapping; gene expression control; promoter; termination sequence.

XX

OS Arabidopsis thaliana.

XX

PN EP1033405-A2.

XX

PD 06-SEP-2000.

XX

PP 25-FEB-2000; 2000EP-0301439.

XX

PR 25-FEB-1999; 99US-0121825.

PR 05-MAR-1999; 99US-0123180.

PR 09-MAR-1999; 99US-0123548.

PR 23-MAR-1999; 99US-0127788.

PR 25-MAR-1999; 99US-0126264.

PR 29-MAR-1999; 99US-0126785.

PR 01-APR-1999; 99US-0127462.

PR 06-APR-1999; 99US-0128234.

PR 08-APR-1999; 99US-0128714.

PR 16-APR-1999; 99US-0128845.

PR 19-APR-1999; 99US-0130077.

PR 21-APR-1999; 99US-0130449.

PR 23-APR-1999; 99US-0130510.

PR 23-APR-1999; 99US-0130891.

PR 28-APR-1999; 99US-0131449.

PR 30-APR-1999; 99US-0132048.

PR 04-MAY-1999; 99US-0132407.

PR 05-MAY-1999; 99US-0132485.

PR 06-MAY-1999; 99US-0132486.

PR 07-MAY-1999; 99US-0132863.

PR 11-MAY-1999; 99US-0134256.

PR 14-MAY-1999; 99US-0134219.

PR 14-MAY-1999; 99US-0134221.

PR 18-MAY-1999; 99US-0134370.

PR 19-MAY-1999; 99US-0134941.

PR 20-MAY-1999; 99US-0135124.

PR 21-MAY-1999; 99US-0135253.

PR 24-MAY-1999; 99US-0135629.

PR 25-MAY-1999; 99US-0136021.

PR	27-MAY-1999;	990US-0136392.	PR	04-AUG-1999;	990US-0147204.
PR	28-MAY-1999;	990US-0136782.	PR	04-AUG-1999;	990US-0147702.
PR	01-JUN-1999;	990US-0137222.	PR	05-AUG-1999;	990US-0147192.
PR	03-JUN-1999;	990US-0137538.	PR	05-AUG-1999;	990US-0147260.
PR	04-JUN-1999;	990US-0137502.	PR	06-AUG-1999;	990US-0147703.
PR	07-JUN-1999;	990US-0137724.	PR	06-AUG-1999;	990US-0147416.
PR	08-JUN-1999;	990US-0138054.	PR	09-AUG-1999;	990US-0147493.
PR	10-JUN-1999;	990US-0138540.	PR	09-AUG-1999;	990US-0147935.
PR	14-JUN-1999;	990US-0139119.	PR	10-AUG-1999;	990US-0148171.
PR	16-JUN-1999;	990US-0139452.	PR	11-AUG-1999;	990US-0148319.
PR	16-JUN-1999;	990US-0139453.	PR	12-AUG-1999;	990US-0148341.
PR	17-JUN-1999;	990US-0139454.	PR	13-AUG-1999;	990US-0148565.
PR	18-JUN-1999;	990US-0139455.	PR	13-AUG-1999;	990US-0148684.
PR	18-JUN-1999;	990US-0139487.	PR	16-AUG-1999;	990US-0149668.
PR	18-JUN-1999;	990US-0139491.	PR	17-AUG-1999;	990US-0149175.
PR	18-JUN-1999;	990US-0139492.	PR	18-AUG-1999;	990US-0149426.
PR	18-JUN-1999;	990US-0139493.	PR	20-AUG-1999;	990US-0149722.
PR	18-JUN-1999;	990US-0139495.	PR	20-AUG-1999;	990US-0149723.
PR	18-JUN-1999;	990US-0139496.	PR	20-AUG-1999;	990US-0149929.
PR	18-JUN-1999;	990US-0139497.	PR	23-AUG-1999;	990US-0149902.
PR	18-JUN-1999;	990US-0139498.	PR	23-AUG-1999;	990US-0149930.
PR	23-JUN-1999;	990US-0140353.	PR	25-AUG-1999;	990US-0150566.
PR	23-JUN-1999;	990US-0140354.	PR	26-AUG-1999;	990US-0150884.
PR	18-JUN-1999;	990US-0139750.	PR	27-AUG-1999;	990US-0151065.
PR	21-JUN-1999;	990US-0139763.	PR	27-AUG-1999;	990US-0151066.
PR	21-JUN-1999;	990US-0139817.	PR	27-AUG-1999;	990US-0151080.
PR	22-JUN-1999;	990US-0139899.	PR	30-AUG-1999;	990US-0151303.
PR	22-JUN-1999;	990US-0140287.	PR	31-AUG-1999;	990US-0151438.
PR	01-JUL-1999;	990US-0140442.	PR	01-SEP-1999;	990US-0151930.
PR	01-JUL-1999;	990US-0142154.	PR	07-SEP-1999;	990US-0152363.
PR	24-JUN-1999;	990US-0140955.	PR	10-SEP-1999;	990US-0153070.
PR	28-JUN-1999;	990US-0140956.	PR	13-SEP-1999;	990US-0153758.
PR	29-JUN-1999;	990US-0140957.	PR	15-SEP-1999;	990US-0154018.
PR	30-JUN-1999;	990US-0140959.	PR	16-SEP-1999;	990US-0154039.
PR	12-JUL-1999;	990US-0142977.	PR	20-SEP-1999;	990US-0154779.
PR	13-JUL-1999;	990US-0143542.	PR	22-SEP-1999;	990US-0155139.
PR	01-JUL-1999;	990US-0142155.	PR	23-SEP-1999;	990US-0155486.
PR	02-JUL-1999;	990US-0142055.	PR	24-SEP-1999;	990US-0155659.
PR	06-JUL-1999;	990US-0142390.	PR	28-SEP-1999;	990US-0156458.
PR	08-JUL-1999;	990US-0142803.	PR	29-SEP-1999;	990US-0156459.
PR	09-JUL-1999;	990US-0142920.	PR	04-OCT-1999;	990US-0157117.
PR	12-JUL-1999;	990US-0142977.	PR	05-OCT-1999;	990US-0157753.
PR	13-JUL-1999;	990US-0143542.	PR	06-OCT-1999;	990US-0157865.
PR	14-JUL-1999;	990US-0143624.	PR	07-OCT-1999;	990US-0158029.
PR	15-JUL-1999;	990US-0144005.	PR	08-OCT-1999;	990US-0158232.
PR	16-JUL-1999;	990US-0144085.	PR	12-OCT-1999;	990US-0158369.
PR	16-JUL-1999;	990US-0144086.	PR	13-OCT-1999;	990US-0158293.
PR	19-JUL-1999;	990US-0144325.	PR	13-OCT-1999;	990US-0158294.
PR	19-JUL-1999;	990US-0144331.	PR	13-OCT-1999;	990US-0159295.
PR	19-JUL-1999;	990US-0144332.	PR	13-OCT-1999;	990US-0159330.
PR	19-JUL-1999;	990US-0144333.	PR	14-OCT-1999;	990US-0159331.
PR	19-JUL-1999;	990US-0144334.	PR	14-OCT-1999;	990US-0159637.
PR	19-JUL-1999;	990US-0144335.	PR	14-OCT-1999;	990US-0159638.
PR	20-JUL-1999;	990US-0144352.	PR	14-OCT-1999;	990US-0159638.
PR	20-JUL-1999;	990US-0144632.	PR	18-OCT-1999;	990US-0159584.
PR	21-JUL-1999;	990US-0144884.	PR	21-OCT-1999;	990US-0160741.
PR	21-JUL-1999;	990US-0144884.	PR	21-OCT-1999;	990US-0160747.
PR	22-JUL-1999;	990US-0145089.	PR	21-OCT-1999;	990US-0160768.
PR	22-JUL-1999;	990US-0145089.	PR	21-OCT-1999;	990US-0160770.
PR	22-JUL-1999;	990US-0145192.	PR	21-OCT-1999;	990US-0160814.
PR	23-JUL-1999;	990US-0145145.	PR	21-OCT-1999;	990US-0160815.
PR	23-JUL-1999;	990US-0145218.	PR	22-OCT-1999;	990US-0160980.
PR	23-JUL-1999;	990US-0145285.	PR	22-OCT-1999;	990US-0160981.
PR	26-JUL-1999;	990US-0145224.	PR	22-OCT-1999;	990US-0161361.
PR	26-JUL-1999;	990US-0145087.	PR	25-OCT-1999;	990US-0161404.
PR	27-JUL-1999;	990US-0145089.	PR	25-OCT-1999;	990US-0161405.
PR	27-JUL-1999;	990US-0145192.	PR	25-OCT-1999;	990US-0161406.
PR	27-JUL-1999;	990US-0145191.	PR	26-OCT-1999;	990US-0161359.
PR	28-JUL-1999;	990US-0145951.	PR	26-OCT-1999;	990US-0161500.
PR	02-AUG-1999;	990US-016386.	PR	26-OCT-1999;	990US-0161361.
PR	02-AUG-1999;	990US-016388.	PR	28-OCT-1999;	990US-0161920.
PR	03-AUG-1999;	990US-016389.	PR	28-OCT-1999;	990US-0161922.
PR	03-AUG-1999;	990US-0147038.	PR	28-OCT-1999;	990US-0161993.

PR	29-OCT-1999;	99US-0162142.	PR	10-JUN-1999;	99US-0138540.
ID	AAC49296	standard; Protein; 287 AA.	PR	10-JUN-1999;	99US-0138847.
XX			PR	14-JUN-1999;	99US-0139119.
AC	AAC49296;		PR	16-JUN-1999;	99US-0139452.
XX			PR	16-JUN-1999;	99US-0139453.
DT	18-OCT-2000 (first entry)		PR	17-JUN-1999;	99US-0139492.
XX			PR	18-JUN-1999;	99US-0139454.
DE	Arabidopsis thaliana protein fragment SEQ ID NO: 62350.		PR	18-JUN-1999;	99US-0139455.
XX			PR	18-JUN-1999;	99US-0139456.
KW	Protein identification; signal transduction pathway; metabolic pathway;		PR	18-JUN-1999;	99US-0139463.
KW	hybridisation assay; genetic mapping; gene expression control; promoter;		PR	18-JUN-1999;	99US-0139750.
KW	termination sequence.		PR	18-JUN-1999;	99US-0139763.
OS	Arabidopsis thaliana.		PR	21-JUN-1999;	99US-0139817.
XX			PR	22-JUN-1999;	99US-0139846.
PN	EP1033405-A2.		PR	18-JUN-1999;	99US-0139461.
XX			PR	18-JUN-1999;	99US-0139462.
PD	06-SEP-2000.		PR	23-JUN-1999;	99US-0140353.
XX			PR	23-JUN-1999;	99US-0140354.
PR	25-FEB-1999;	99US-0121825.	PR	24-JUN-1999;	99US-0140695.
PR	05-MAR-1999;	99US-0123180.	PR	28-JUN-1999;	99US-0140823.
PR	09-MAR-1999;	99US-0123548.	PR	29-JUN-1999;	99US-0140991.
PR	23-MAR-1999;	99US-0125788.	PR	30-JUN-1999;	99US-0141287.
PR	25-MAR-1999;	99US-0126264.	PR	01-JUL-1999;	99US-0141842.
PR	29-MAR-1999;	99US-0126785.	PR	01-JUL-1999;	99US-0142154.
PR	01-APR-1999;	99US-0124462.	PR	02-JUL-1999;	99US-0142055.
PR	06-APR-1999;	99US-0128234.	PR	06-JUL-1999;	99US-0142390.
PR	08-APR-1999;	99US-0128714.	PR	08-JUL-1999;	99US-0142803.
PR	16-APR-1999;	99US-0128845.	PR	09-JUL-1999;	99US-0142920.
PR	19-APR-1999;	99US-0130077.	PR	12-JUL-1999;	99US-014297.
PR	21-APR-1999;	99US-0130449.	PR	13-JUL-1999;	99US-0143542.
PR	23-APR-1999;	99US-0130510.	PR	14-JUL-1999;	99US-0143624.
PR	23-APR-1999;	99US-0130891.	PR	15-JUL-1999;	99US-0144005.
PR	28-APR-1999;	99US-0131449.	PR	16-JUL-1999;	99US-0144085.
PR	30-APR-1999;	99US-0132048.	PR	17-JUL-1999;	99US-0144125.
PR	19-MAY-1999;	99US-0132484.	PR	19-JUL-1999;	99US-0144331.
PR	04-MAY-1999;	99US-0132485.	PR	19-JUL-1999;	99US-0144332.
PR	05-MAY-1999;	99US-0132486.	PR	19-JUL-1999;	99US-0144884.
PR	06-MAY-1999;	99US-0132487.	PR	19-JUL-1999;	99US-0144914.
PR	07-MAY-1999;	99US-0132863.	PR	21-JUL-1999;	99US-0144915.
PR	11-MAY-1999;	99US-0134256.	PR	21-JUL-1999;	99US-0145086.
PR	14-MAY-1999;	99US-0134218.	PR	22-JUL-1999;	99US-0145085.
PR	14-MAY-1999;	99US-0134219.	PR	22-JUL-1999;	99US-0145087.
PR	14-MAY-1999;	99US-0134221.	PR	22-JUL-1999;	99US-0145089.
PR	14-MAY-1999;	99US-0134221.	PR	21-JUL-1999;	99US-0144814.
PR	14-MAY-1999;	99US-0134221.	PR	21-JUL-1999;	99US-0144334.
PR	18-MAY-1999;	99US-0134270.	PR	19-JUL-1999;	99US-0144335.
PR	19-MAY-1999;	99US-0134768.	PR	20-JUL-1999;	99US-0144352.
PR	20-MAY-1999;	99US-0135124.	PR	20-JUL-1999;	99US-0144632.
PR	21-MAY-1999;	99US-0135253.	PR	20-JUL-1999;	99US-0144884.
PR	24-MAY-1999;	99US-0136229.	PR	21-JUL-1999;	99US-0144914.
PR	25-MAY-1999;	99US-0136021.	PR	21-JUL-1999;	99US-0144915.
PR	27-MAY-1999;	99US-013692.	PR	23-JUL-1999;	99US-0145218.
PR	28-MAY-1999;	99US-0136782.	PR	23-JUL-1999;	99US-0145224.
PR	01-JUN-1999;	99US-0137222.	PR	26-JUL-1999;	99US-014526.
PR	03-JUN-1999;	99US-0137528.	PR	27-JUL-1999;	99US-0145913.
PR	04-JUN-1999;	99US-0137502.	PR	27-JUL-1999;	99US-0145918.
PR	07-JUN-1999;	99US-0137224.	PR	27-JUL-1999;	99US-0147038.
PR	08-JUN-1999;	99US-0138094.	PR	04-AUG-1999;	99US-0147204.
PR	09-AUG-1999;	99US-0147493.	PR	04-AUG-1999;	99US-0147302.

PR	09-AUG-1999;	99US-0147935.	Db	
PR	11-AUG-1999;	99US-0118171.	229	dif9wgk 234
PR	12-AUG-1999;	99US-0148341.	RESULT	5
PR	13-AUG-1999;	99US-018565.	ID	AAG07708
PR	13-AUG-1999;	99US-0148684.	standard:	protein: 293 AA.
PR	16-AUG-1999;	99US-0149368.	AC	AAG07708;
PR	17-AUG-1999;	99US-019175.	XX	
PR	18-AUG-1999;	99US-014426.	XX	
PR	20-AUG-1999;	99US-0149723.	XX	
PR	20-AUG-1999;	99US-014929.	DT	17-OCT-2000 (first entry)
PR	20-AUG-1999;	99US-0149902.	XX	
PR	23-AUG-1999;	99US-0149930.	DE	Arabidopsis thaliana protein fragment SEQ ID NO: 4963.
PR	25-AUG-1999;	99US-0150566.	XX	
PR	26-AUG-1999;	99US-0150884.	KW	Protein identification; signal transduction pathway; metabolic pathway;
PR	27-AUG-1999;	99US-0151065.	KW	hybridisation assay; genetic mapping; gene expression control; promoter;
PR	27-AUG-1999;	99US-0151066.	KW	termination sequence.
PR	27-AUG-1999;	99US-0151080.	XX	
PR	30-AUG-1999;	99US-0151303.	OS	Arabidopsis thaliana.
PR	31-AUG-1999;	99US-0151348.	XX	
PR	01-SEP-1999;	99US-0151930.	PN	
PR	07-SEP-1999;	99US-0152363.	XX	
PR	10-SEP-1999;	99US-0153070.	PD	06-SEP-2000.
PR	13-SEP-1999;	99US-0153758.	XX	
PR	15-SEP-1999;	99US-0154018.	PP	25-FEB-2000; 2000EP-0301439.
PR	16-SEP-1999;	99US-0154039.	XX	
PR	20-SEP-1999;	99US-0154779.	PR	25-FEB-1999; 99US-0121825.
PR	22-SEP-1999;	99US-0155139.	PR	05-MAR-1999; 99US-0123180.
PR	23-SEP-1999;	99US-0155486.	PR	09-MAR-1999; 99US-0123348.
PR	24-SEP-1999;	99US-0155659.	PR	23-MAR-1999; 99US-0125788.
PR	28-SEP-1999;	99US-0156458.	PR	25-MAR-1999; 99US-0126284.
PR	29-SEP-1999;	99US-0156596.	PR	29-MAR-1999; 99US-0126785.
PR	01-OCT-1999;	99US-0157117.	PR	01-APR-1999; 99US-0127662.
PR	03-OCT-1999;	99US-0157753.	PR	06-APR-1999; 99US-0128234.
PR	05-OCT-1999;	99US-0157865.	PR	08-APR-1999; 99US-0128714.
PR	07-OCT-1999;	99US-0158029.	PR	16-APR-1999; 99US-0129845.
PR	08-OCT-1999;	99US-0158232.	PR	19-APR-1999; 99US-0130077.
PR	12-OCT-1999;	99US-0158369.	PR	21-APR-1999; 99US-0130449.
PR	13-OCT-1999;	99US-0159293.	PR	23-APR-1999; 99US-0130510.
PR	14-OCT-1999;	99US-0159638.	PR	23-APR-1999; 99US-0130591.
PR	13-OCT-1999;	99US-0159294.	PR	28-APR-1999; 99US-0131449.
PR	14-OCT-1999;	99US-0159295.	PR	30-APR-1999; 99US-0132048.
PR	14-OCT-1999;	99US-0159329.	PR	19-APR-1999; 99US-0132407.
PR	14-OCT-1999;	99US-0159330.	PR	01-MAY-1999; 99US-0132884.
PR	14-OCT-1999;	99US-0159331.	PR	05-MAY-1999; 99US-0132485.
PR	14-OCT-1999;	99US-0159337.	PR	06-MAY-1999; 99US-0132486.
PR	21-OCT-1999;	99US-0160707.	PR	06-MAY-1999; 99US-0132487.
PR	21-OCT-1999;	99US-0160814.	PR	07-MAY-1999; 99US-0132663.
PR	21-OCT-1999;	99US-0160815.	PR	11-MAY-1999; 99US-01324256.
PR	22-OCT-1999;	99US-0160741.	PR	14-MAY-1999; 99US-0134218.
PR	21-OCT-1999;	99US-0160767.	PR	14-MAY-1999; 99US-0134219.
PR	21-OCT-1999;	99US-0160770.	PR	14-MAY-1999; 99US-0134221.
PR	21-OCT-1999;	99US-0160814.	PR	14-MAY-1999; 99US-0134370.
PR	21-OCT-1999;	99US-0160815.	PR	18-MAY-1999; 99US-0134768.
PR	22-OCT-1999;	99US-0160801.	PR	19-MAY-1999; 99US-0134941.
PR	22-OCT-1999;	99US-0160981.	PR	20-MAY-1999; 99US-0135124.
PR	26-OCT-1999;	99US-0161361.	PR	21-MAY-1999; 99US-0135253.
PR	25-OCT-1999;	99US-0161404.	PR	24-MAY-1999; 99US-0135259.
PR	25-OCT-1999;	99US-0161405.	PR	25-MAY-1999; 99US-0136201.
PR	25-OCT-1999;	99US-0161406.	PR	27-MAY-1999; 99US-0136592.
PR	26-OCT-1999;	99US-0161359.	PR	28-MAY-1999; 99US-0136782.
PR	26-OCT-1999;	99US-0161360.	PR	01-JUN-1999; 99US-0137222.
PR	28-OCT-1999;	99US-0161920.	PR	03-JUN-1999; 99US-0137528.
Matches	6;	Conservative	0;	Mismatches
1 DFGWGK 6				

PR	18-JUN-1999;	99US-0139455.	PR	17-AUG-1999;	99US-0149175.
PR	18-JUN-1999;	99US-0139456.	PR	18-AUG-1999;	99US-0149426.
PR	18-JUN-1999;	99US-0139457.	PR	20-AUG-1999;	99US-0149723.
PR	18-JUN-1999;	99US-0139458.	PR	20-AUG-1999;	99US-0149929.
PR	18-JUN-1999;	99US-0139459.	PR	23-AUG-1999;	99US-0149902.
PR	18-JUN-1999;	99US-0139460.	PR	23-AUG-1999;	99US-0149930.
PR	18-JUN-1999;	99US-0139461.	PR	25-AUG-1999;	99US-0150566.
PR	18-JUN-1999;	99US-0139462.	PR	26-AUG-1999;	99US-0150884.
PR	18-JUN-1999;	99US-0139463.	PR	27-AUG-1999;	99US-0151065.
PR	18-JUN-1999;	99US-0139750.	PR	27-AUG-1999;	99US-0151066.
PR	18-JUN-1999;	99US-0139763.	PR	27-AUG-1999;	99US-0151080.
PR	21-JUN-1999;	99US-0139817.	PR	30-AUG-1999;	99US-0151303.
PR	22-JUN-1999;	99US-0139899.	PR	31-AUG-1999;	99US-0151438.
PR	23-JUN-1999;	99US-0140353.	PR	01-SEP-1999;	99US-0151930.
PR	23-JUN-1999;	99US-0140354.	PR	07-SEP-1999;	99US-0152363.
PR	24-JUN-1999;	99US-0140695.	PR	10-SEP-1999;	99US-0153070.
PR	28-JUN-1999;	99US-0140823.	PR	13-SEP-1999;	99US-0153758.
PR	29-JUN-1999;	99US-0140991.	PR	15-SEP-1999;	99US-0154018.
PR	30-JUN-1999;	99US-0141287.	PR	16-SEP-1999;	99US-0154039.
PR	30-JUN-1999;	99US-0141842.	PR	20-SEP-1999;	99US-0151930.
PR	01-JUL-1999;	99US-0142154.	PR	22-SEP-1999;	99US-0152363.
PR	02-JUL-1999;	99US-0142055.	PR	23-SEP-1999;	99US-0153070.
PR	06-JUL-1999;	99US-0142390.	PR	24-SEP-1999;	99US-0153758.
PR	08-JUL-1999;	99US-0142803.	PR	28-SEP-1999;	99US-0154658.
PR	09-JUL-1999;	99US-0143920.	PR	29-SEP-1999;	99US-0156596.
PR	12-JUL-1999;	99US-0142977.	PR	04-OCT-1999;	99US-0157117.
PR	13-JUL-1999;	99US-0143542.	PR	05-OCT-1999;	99US-0157753.
PR	14-JUL-1999;	99US-0143624.	PR	06-OCT-1999;	99US-0157865.
PR	15-JUL-1999;	99US-0144005.	PR	07-OCT-1999;	99US-0158029.
PR	16-JUL-1999;	99US-0144085.	PR	08-OCT-1999;	99US-0158369.
PR	17-JUL-1999;	99US-0144086.	PR	12-OCT-1999;	99US-0158369.
PR	19-JUL-1999;	99US-0144325.	PR	13-OCT-1999;	99US-0159293.
PR	19-JUL-1999;	99US-0144331.	PR	13-OCT-1999;	99US-0159294.
PR	19-JUL-1999;	99US-0144332.	PR	13-OCT-1999;	99US-0159295.
PR	19-JUL-1999;	99US-0144333.	PR	14-OCT-1999;	99US-0159329.
PR	19-JUL-1999;	99US-0144334.	PR	14-OCT-1999;	99US-0159330.
PR	19-JUL-1999;	99US-0144335.	PR	14-OCT-1999;	99US-0159331.
PR	19-JUL-1999;	99US-0144352.	PR	14-OCT-1999;	99US-0159637.
PR	19-JUL-1999;	99US-0144384.	PR	14-OCT-1999;	99US-0159638.
PR	20-JUL-1999;	99US-0144814.	PR	18-OCT-1999;	99US-0159588.
PR	21-JUL-1999;	99US-0144815.	PR	21-OCT-1999;	99US-0160741.
PR	21-JUL-1999;	99US-0145086.	PR	21-OCT-1999;	99US-0160767.
PR	21-JUL-1999;	99US-0145088.	PR	21-OCT-1999;	99US-0160768.
PR	22-JUL-1999;	99US-0145087.	PR	21-OCT-1999;	99US-0160770.
PR	22-JUL-1999;	99US-0145089.	PR	21-OCT-1999;	99US-0160814.
PR	22-JUL-1999;	99US-0145090.	PR	21-OCT-1999;	99US-0160815.
PR	23-JUL-1999;	99US-0145145.	PR	22-OCT-1999;	99US-0160980.
PR	23-JUL-1999;	99US-0145218.	PR	22-OCT-1999;	99US-0160981.
PR	23-JUL-1999;	99US-0145224.	PR	22-OCT-1999;	99US-0160989.
PR	26-JUL-1999;	99US-0145276.	PR	25-OCT-1999;	99US-0161404.
PR	27-JUL-1999;	99US-0145913.	PR	25-OCT-1999;	99US-0161405.
PR	27-JUL-1999;	99US-0145918.	PR	25-OCT-1999;	99US-0161406.
PR	27-JUL-1999;	99US-0145919.	PR	26-OCT-1999;	99US-0161359.
PR	28-JUL-1999;	99US-0145951.	PR	26-OCT-1999;	99US-0161360.
PR	02-AUG-1999;	99US-0146386.	PR	26-OCT-1999;	99US-0161361.
PR	02-AUG-1999;	99US-0146388.	PR	28-OCT-1999;	99US-0161920.
PR	02-AUG-1999;	99US-0146389.	PR	28-OCT-1999;	99US-0161922.
PR	03-AUG-1999;	99US-0147038.	PR	28-OCT-1999;	99US-0161933.
PR	04-AUG-1999;	99US-0147204.	PR	29-OCT-1999;	99US-0162142.
PR	04-AUG-1999;	99US-0147302.			
PR	05-AUG-1999;	99US-0148171.			
PR	11-AUG-1999;	99US-0148319.			
PR	12-AUG-1999;	99US-0148365.			
PR	13-AUG-1999;	99US-0148684.			
PR	16-AUG-1999;	99US-0149368.			

RESULT 6  
ID AAG4925  
AAG4925 standard; Protein; 293 AA.  
Qy 1 DPGWKG 6  
Oy 1 DPGWKG 6  
Db 235 dfgwgk 240

Query Match Score 100.0%; DB 21; Length 293;  
Best Local Similarity 100.0%; Pred. No. 26; Mismatches 0; Indels 0; Gaps 0;

Matches 6; Conservative 0; Misnatches 0; Indels 0; Gaps 0;

✓ Mon Dec 10 13:33:27 2001

XX		PR	18-JUN-1999;	99US-0139462.
AC		PR	18-JUN-1999;	99US-0139433.
XX		PR	18-JUN-1999;	99US-0139755.
DT		PR	18-JUN-1999;	99US-0139763.
XX		PR	21-JUN-1999;	99US-0139817.
DE		PR	22-JUL-1999;	99US-0139899.
XX		PR	23-JUN-1999;	99US-0140353.
KW	Arabidopsis thaliana protein fragment SEQ ID NO: 62349.	PR	23-JUN-1999;	99US-0140354.
KW	protein identification; signal transduction pathway; metabolic pathway;	PR	24-JUL-1999;	99US-0140695.
KW	hybridisation assay; genetic mapping; gene expression control; promoter;	PR	28-JUN-1999;	99US-0140823.
XX	termination sequence.	PR	29-JUN-1999;	99US-0140931.
OS	Arabidopsis thaliana.	PR	30-JUN-1999;	99US-0142887.
XX		PR	01-JUL-1999;	99US-0141842.
PN	EP1033405-A2.	PR	01-JUL-1999;	99US-0142154.
XX		PR	02-JUL-1999;	99US-0142055.
PD	06-SEP-2000.	PR	06-JUL-1999;	99US-0142390.
XX		PR	08-JUL-1999;	99US-0142803.
PF	25-FEB-2000; 2000EP-0301439.	PR	09-JUL-1999;	99US-0142220.
XX		PR	12-JUL-1999;	99US-0142377.
PR	25-FEB-1999; 99US-0121825.	PR	13-JUL-1999;	99US-0143342.
PR	05-MAR-1999; 99US-0123180.	PR	14-JUL-1999;	99US-0143624.
PR	09-MAR-1999; 99US-0123548.	PR	15-JUL-1999;	99US-0144005.
PR	23-MAR-1999; 99US-0125788.	PR	16-JUL-1999;	99US-0144085.
PR	25-MAR-1999; 99US-0126254.	PR	16-JUL-1999;	99US-0144287.
PR	29-MAR-1999; 99US-0126785.	PR	19-JUL-1999;	99US-0142325.
PR	01-APR-1999; 99US-0127462.	PR	19-JUL-1999;	99US-014331.
PR	06-APR-1999; 99US-0128334.	PR	19-JUL-1999;	99US-014332.
PR	08-APR-1999; 99US-0128714.	PR	19-JUL-1999;	99US-0144333.
PR	16-APR-1999; 99US-0129845.	PR	19-JUL-1999;	99US-0144334.
PR	19-APR-1999; 99US-0130077.	PR	19-JUL-1999;	99US-0144335.
PR	21-APR-1999; 99US-0130449.	PR	20-JUL-1999;	99US-0144352.
PR	23-APR-1999; 99US-0130510.	PR	20-JUL-1999;	99US-0144632.
PR	23-APR-1999; 99US-0130591.	PR	20-JUL-1999;	99US-0144884.
PR	28-APR-1999; 99US-0131449.	PR	21-JUL-1999;	99US-0144814.
PR	30-APR-1999; 99US-0132008.	PR	23-JUL-1999;	99US-0145145.
PR	30-APR-1999; 99US-0132407.	PR	21-JUL-1999;	99US-0145086.
PR	04-MAY-1999; 99US-0132484.	PR	21-JUL-1999;	99US-0145335.
PR	05-MAY-1999; 99US-0132485.	PR	22-JUL-1999;	99US-0144352.
PR	06-MAY-1999; 99US-0132486.	PR	22-JUL-1999;	99US-0145087.
PR	06-MAY-1999; 99US-0132487.	PR	22-JUL-1999;	99US-0145089.
PR	07-MAY-1999; 99US-0132833.	PR	22-JUL-1999;	99US-0145192.
PR	11-MAY-1999; 99US-0132408.	PR	22-JUL-1999;	99US-0145199.
PR	14-MAY-1999; 99US-0132484.	PR	23-JUL-1999;	99US-0145218.
PR	14-MAY-1999; 99US-0132419.	PR	23-JUL-1999;	99US-0145224.
PR	14-MAY-1999; 99US-0134221.	PR	26-JUL-1999;	99US-0145276.
PR	14-MAY-1999; 99US-0134370.	PR	27-JUL-1999;	99US-0145913.
PR	18-MAY-1999; 99US-0134370.	PR	27-JUL-1999;	99US-0145918.
PR	27-MAY-1999; 99US-0136021.	PR	27-JUL-1999;	99US-0145918.
PR	28-MAY-1999; 99US-0136592.	PR	28-JUL-1999;	99US-0145919.
PR	01-JUN-1999; 99US-0134241.	PR	02-AUG-1999;	99US-0145951.
PR	03-JUN-1999; 99US-0135533.	PR	02-AUG-1999;	99US-0146386.
PR	24-MAY-1999; 99US-0135629.	PR	02-AUG-1999;	99US-0146388.
PR	25-MAY-1999; 99US-0136021.	PR	03-AUG-1999;	99US-0147038.
PR	27-MAY-1999; 99US-0136021.	PR	04-AUG-1999;	99US-0147204.
PR	28-MAY-1999; 99US-0136782.	PR	04-AUG-1999;	99US-0147302.
PR	01-JUN-1999; 99US-0137222.	PR	05-AUG-1999;	99US-0147192.
PR	14-JUN-1999; 99US-0137528.	PR	05-AUG-1999;	99US-0148319.
PR	04-JUN-1999; 99US-0137502.	PR	06-AUG-1999;	99US-0147303.
PR	07-JUN-1999; 99US-0137724.	PR	06-AUG-1999;	99US-0147416.
PR	08-JUN-1999; 99US-0138094.	PR	09-AUG-1999;	99US-0147493.
PR	10-JUN-1999; 99US-0138840.	PR	09-AUG-1999;	99US-0147935.
PR	14-JUN-1999; 99US-0138847.	PR	10-AUG-1999;	99US-0148171.
PR	16-JUN-1999; 99US-0139452.	PR	11-AUG-1999;	99US-0148260.
PR	18-JUN-1999; 99US-0139453.	PR	12-AUG-1999;	99US-014831.
PR	17-JUN-1999; 99US-0139492.	PR	13-AUG-1999;	99US-0148565.
PR	18-JUN-1999; 99US-0139454.	PR	16-AUG-1999;	99US-0149368.
PR	18-JUN-1999; 99US-0139455.	PR	17-AUG-1999;	99US-0149175.
PR	18-JUN-1999; 99US-0139456.	PR	18-AUG-1999;	99US-0149426.
PR	18-JUN-1999; 99US-0139457.	PR	20-AUG-1999;	99US-0149722.
PR	18-JUN-1999; 99US-0139459.	PR	20-AUG-1999;	99US-0149723.
PR	18-JUN-1999; 99US-0139461.	PR	23-AUG-1999;	99US-0149930.

PR 25-AUG-1999; 99US-0150566.  
 PR 26-AUG-1999; 99US-0150884.  
 PR 27-AUG-1999; 99US-0151065.  
 PR 27-AUG-1999; 99US-0151066.  
 PR 30-AUG-1999; 99US-0151080.  
 PR 31-AUG-1999; 99US-0151303.  
 PR 01-SEP-1999; 99US-0151930.  
 PR 07-SEP-1999; 99US-0152363.  
 PR 10-SEP-1999; 99US-0153070.  
 PR 13-SEP-1999; 99US-0153758.  
 PR 15-SEP-1999; 99US-0154018.  
 PR 16-SEP-1999; 99US-0154039.  
 PR 20-SEP-1999; 99US-0154779.  
 PR 22-SEP-1999; 99US-0155139.  
 PR 23-SEP-1999; 99US-0155486.  
 PR 24-SEP-1999; 99US-0155659.  
 PR 28-SEP-1999; 99US-0156458.  
 PR 29-SEP-1999; 99US-0156596.  
 PR 04-OCT-1999; 99US-0157117.  
 PR 05-OCT-1999; 99US-0157753.  
 PR 06-OCT-1999; 99US-0157865.  
 PR 07-OCT-1999; 99US-0158029.  
 PR 08-OCT-1999; 99US-0158232.  
 PR 12-OCT-1999; 99US-0158369.  
 PR 13-OCT-1999; 99US-0158293.  
 PR 13-OCT-1999; 99US-0158294.  
 PR 14-OCT-1999; 99US-0158295.  
 PR 14-OCT-1999; 99US-0158329.  
 PR 14-OCT-1999; 99US-0158330.  
 PR 14-OCT-1999; 99US-0158331.  
 PR 14-OCT-1999; 99US-0158637.  
 PR 14-OCT-1999; 99US-0158638.  
 PR 18-OCT-1999; 99US-0158584.  
 PR 21-OCT-1999; 99US-0160741.  
 PR 21-OCT-1999; 99US-0160767.  
 PR 21-OCT-1999; 99US-0160768.  
 PR 21-OCT-1999; 99US-0160770.  
 PR 21-OCT-1999; 99US-0160814.  
 PR 21-OCT-1999; 99US-0160815.  
 PR 22-OCT-1999; 99US-0160980.  
 PR 22-OCT-1999; 99US-0160981.  
 PR 22-OCT-1999; 99US-0160989.  
 PR 25-OCT-1999; 99US-0161404.  
 PR 25-OCT-1999; 99US-0161406.  
 PR 26-OCT-1999; 99US-0161359.  
 PR 26-OCT-1999; 99US-0161360.  
 PR 26-OCT-1999; 99US-0161361.  
 PR 28-OCT-1999; 99US-0161920.  
 PR 28-OCT-1999; 99US-0161992.  
 PR 28-OCT-1999; 99US-0161993.  
 PR 29-OCT-1999; 99US-0162142.  
  
 Query Match 100.0%; Score 40; DB 21; Length 293;  
 Best Local Similarity 100.0%; Pred. No. 26; Mismatches 0;  
 Matches 6; Conservative 0; Indels 0; Gaps 0;  
 Oy 1 DFGWKG 6  
 Db 297 dfgwkg 302  
 Dg t  
  
 RESULT 7  
 AAE00230 standard; Protein: 302 AA.  
 ID AAE00230  
 AC AAE00230;  
 XX  
 DT 13-JUN-2001 (first entry)  
 XX  
 DE Protein encoded by Probe 4.  
 XX  
  
 RESULT 7  
 AAE00230 standard; Protein: 302 AA.  
 ID AAE00230  
 AC AAE00230;  
 XX  
 DT 13-JUN-2001 (first entry)  
 XX  
 DE Protein encoded by Probe 5.  
 XX  
  
 RESULT 8  
 AAE00231 standard; Protein: 302 AA.  
 ID AAE00231  
 AC AAE00231;  
 XX  
 DT 13-JUN-2001 (first entry)  
 XX  
 DE Protein encoded by Probe 5.  
 XX  
  
 RESULT 8  
 AAE00231 standard; Protein: 302 AA.  
 ID AAE00231  
 AC AAE00231;  
 XX  
 DT 13-JUN-2001 (first entry)  
 XX  
 DE Protein encoded by Probe 5.  
 XX  
  
 KW Transacylase; taxol; paclitaxel biosynthesis; taxoid; probe;  
 KW transgenic organism.  
 XX  
 OS Taxus  
 XX cuspidata.  
 XX  
 PN WO200123586-A2.  
 XX  
 PD 05-APR-2001.  
 XX  
 PD ..  
 XX  
 PF 29-SEP-2000; 2000WO-US27006.  
 XX  
 PN WO200123586-A2.  
 XX  
 PR 30-SEP-1999; 99US-0411145.  
 XX  
 PR 07-DEC-1999; 99US-0457046.  
 XX  
 PA (UNIW ) UNIV WASHINGTON STATE RES FOUND.  
 XX  
 PI Croteau RB, Walker KD, Schoendorf A, Wildung MR;  
 XX  
 DR WPI; 2001-245004/25.  
 XX  
 N-PDBB; AAD03336.  
 XX  
 New transacylase enzymes, useful for the high yield production of  
 PT Taxol™, related taxoids and useful intermediates in the  
 PT paclitaxel biosynthetic pathway -  
 PT  
 XX  
 PS Claim 1; Page 101-102; 162pp; English.  
 XX  
 CC The present sequence is the amino acid sequence of probe 4.  
 CC is derived from AT-FOR3 and AT-REV1 primers and is used for screening  
 CC taxus cuspidata transacylase cDNA. The probes are useful for the  
 CC identification of (nucleic acid sequences encoding) transacylases.  
 CC The probes isolated from the *Taxus* genus are useful for the  
 CC synthetic production of Taxol™ and related taxoids, as well as  
 CC intermediates in the paclitaxel biosynthetic pathway. They can also  
 CC be used for the creation of transgenic organisms that either produce  
 CC the transacylases for subsequent in vitro use, or produce the  
 CC transacylases in vivo. The (nucleic acids encoding) transacylases are  
 CC also useful for creating specific binding agents that recognise the  
 CC corresponding transacylases. Binding agents include (fragments of)  
 CC antibodies or any other agent capable of specifically binding to the  
 CC groups on the proteins.  
 XX  
 SQ Sequence 302 AA;  
  
 Query Match 100.0%; Score 40; DB 22; Length 302;  
 Best Local Similarity 100.0%; Pred. No. 27; Mismatches 0;  
 Matches 6; Conservative 0; Indels 0; Gaps 0;  
 Oy 1 DFGWKG 6  
 Db 297 dfgwkg 302  
 Dg t  
  
 RESULT 8  
 AAE00231 standard; Protein: 302 AA.  
 ID AAE00231  
 AC AAE00231;  
 XX  
 DT 13-JUN-2001 (first entry)  
 XX  
 DE Protein encoded by Probe 5.  
 XX  
  
 KW Transacylase; taxol; paclitaxel biosynthesis; taxoid; probe;  
 KW transgenic organism; TAXS.  
 XX  
 OS Taxus  
 XX cuspidata.  
 XX  
 PN WO200123586-A2.  
 XX  
 PD 05-APR-2001.  
 XX



DR WPI; 2001-245004/25.  
 DR N-PSDB; AAD03341.  
 XX  
 PT New transacylase enzymes, useful for the high yield production of  
 PT Taxol(TM), related taxoids and useful intermediates in the in the  
 PT paclitaxel biosynthetic pathway -  
 XX  
 PS Claim 1; Page 108-109; 162pp; English.  
 XX  
 CC The present sequence is the amino acid sequence of probe 9.  
 CC Probe 9 is derived from AT-ROR2 and AT-REV1 primers and is used for screening  
 CC used for screening Taxus cuspidata TAX12 full length cDNA clone. The probes are useful for  
 CC the identification of (nucleic acid sequences encoding) transacylases.  
 CC The probes isolated from the Taxus genus are useful for the  
 CC synthetic production of Taxol(TM) and related taxoids, as well as  
 CC intermediates in the paclitaxel biosynthetic pathway. They can also  
 CC be used for the creation of transgenic organisms that either  
 CC produce the transacylases for subsequent in vitro use, or produce the  
 CC corresponding transacylases. The probes isolated from the Taxus genus  
 CC are useful for the synthetic production of Taxol(TM) and related taxoids,  
 CC as well as intermediates in the paclitaxel biosynthetic pathway. They  
 CC can also be used for the creation of transgenic organisms that either  
 CC produce the transacylases for subsequent in vitro use, or produce the  
 CC transacylases *in vivo*. The (nucleic acids encoding) transacylases are  
 CC also useful for creating specific binding agents that recognise the  
 CC corresponding transacylases. Binding agents include (fragments of)  
 CC antibodies or any other agent capable of specifically binding to the  
 CC groups on the proteins.  
 XX  
 SQ Sequence 302 AA;  
 Query Match 100.0%; Score 40; DB 22; Length 302;  
 Best Local Similarity 100.0%; Pred. No. 27; Mismatches 0;  
 Matches 6; Conservative 0; MisMatches 0; Indels 0; Gaps 0;  
 Qy 1 DFGWCK 6  
 Db 297 dfgwck 302  
 RESULT 11  
 AAE00238  
 ID AAE00238 standard; Protein; 302 AA.  
 XX  
 AC AAE00238;  
 XX  
 DT 13-JUN-2001 (first entry)  
 XX  
 DE Protein encoded by Probe 12.  
 XX  
 KW Transacylase; taxol; paclitaxel biosynthesis; taxoid; probe;  
 KW transgenic organism; TAX12.  
 OS Taxus cuspidata.  
 XX  
 PN WO200123586-A2.  
 XX  
 PD 05-APR-2001.  
 XX  
 PR 29-SEP-2000; 2000WO-US27006.  
 PR 30-SEP-1999; 99US-0411145.  
 PR 07-DEC-1999; 99US-0457046.  
 XX  
 PA (UNIW ) UNIV WASHINGTON STATE RES FOUND.  
 PI Croteau RB, Walker KD, Schoendorf A, Wildung MR;  
 XX  
 PR WPI; 2001-245004/25.  
 PR N-PSDB; AAD03338.  
 XX  
 PT New transacylase enzymes, useful for the high yield production of  
 PT Taxol(TM), related taxoids and useful intermediates in the in the  
 PT paclitaxel biosynthetic pathway -  
 XX  
 PA (UNIW ) UNIV WASHINGTON STATE RES FOUND.  
 PI Croteau RB, Walker KD, Schoendorf A, Wildung MR;  
 XX  
 DR WPI; 2001-245004/25.  
 DR N-PSDB; AAD03344.  
 XX  
 PT New transacylase enzymes, useful for the high yield production of  
 PT Taxol(TM), related taxoids and useful intermediates in the in the  
 PT paclitaxel biosynthetic pathway -  
 XX  
 PS Claim 1; Page 113; 162pp; English.  
 XX  
 CC The present sequence is the amino acid sequence of probe 12. Probe 12  
 CC is derived from AT-ROR3 and AT-REV1 primers and is used for screening  
 CC Taxus cuspidata TAX12 full length cDNA clone. The probes are useful for  
 CC the identification of (nucleic acid sequences encoding) transacylases.  
 CC The probes isolated from the Taxus genus are useful for the  
 CC synthetic production of Taxol(TM) and related taxoids, as well as  
 CC intermediates in the paclitaxel biosynthetic pathway. They can also  
 CC be used for the creation of transgenic organisms that either  
 CC produce the transacylases for subsequent in vitro use, or produce the  
 CC corresponding transacylases *in vivo*. The (nucleic acids encoding) transacylases are  
 CC also useful for creating specific binding agents that recognise the  
 CC corresponding transacylases. Binding agents include (fragments of)  
 CC antibodies or any other agent capable of specifically binding to the  
 CC groups on the proteins.  
 XX  
 SQ Sequence 302 AA;  
 Query Match 100.0%; Score 40; DB 22; Length 302;  
 Best Local Similarity 100.0%; Pred. No. 27; Mismatches 0;  
 Matches 6; Conservative 0; MisMatches 0; Indels 0; Gaps 0;  
 Qy 1 DFGWCK 6  
 Db 297 dfgwck 302  
 RESULT 12  
 AAE00232  
 ID AAE00232 standard; Protein; 303 AA.  
 XX  
 AC AAE00232;  
 XX  
 DT 13-JUN-2001 (first entry)  
 XX  
 DE Protein encoded by Probe 6.  
 XX  
 KW Transacylase; taxol; paclitaxel biosynthesis; transgenic organism; TAX6;  
 KW (10-deacetylbaaccatin III-0-O-acetyltransferase); taxoid; probe.  
 OS Taxus cuspidata.  
 XX  
 PN WO200123586-A2.  
 XX  
 PD 05-APR-2001.  
 XX  
 PR 29-SEP-2000; 2000WO-US27006.  
 PR 30-SEP-1999; 99US-0411145.  
 PR 07-DEC-1999; 99US-0457046.  
 XX  
 PA (UNIW ) UNIV WASHINGTON STATE RES FOUND.  
 PI Croteau RB, Walker KD, Schoendorf A, Wildung MR;  
 XX  
 PR WPI; 2001-245004/25.  
 PR N-PSDB; AAD03338.  
 XX  
 PT New transacylase enzymes, useful for the high yield production of  
 PT Taxol(TM), related taxoids and useful intermediates in the in the  
 PT paclitaxel biosynthetic pathway -  
 XX  
 PS Claim 1; Page 104-105; 162pp; English.  
 XX  
 CC The present sequence is the amino acid sequence of probe 6.  
 CC Probe 6 is derived from AT-ROR2 and AT-REV1 primers and is used for screening  
 CC Taxus cuspidata full length cDNA clone.  
 CC probes are useful for the identification of (nucleic acid sequences  
 CC encoding) transacylases. The probes isolated from the Taxus genus  
 CC are useful for the synthetic production of Taxol(TM) and related taxoids,  
 CC as well as intermediates in the paclitaxel biosynthetic pathway. They  
 CC can also be used for the creation of transgenic organisms that either

CC produce the transacyclases for subsequent in vitro use, or produce the  
 CC transacyclases in vivo. The (nucleic acids encoding) transacyclases are  
 CC also useful for creating specific binding agents that recognise the  
 CC corresponding transacyclases. Binding agents include (fragments of)  
 CC antibodies or any other agent capable of specifically binding to the  
 XX groups on the proteins.

SQ Sequence 303 AA:

Query Match 100.0%; Score 40; DB 22; Length 303;  
 Best Local Similarity 100.0%; Pred. No. 27; Mismatches 0; Indels 0; Gaps 0;  
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 DFGWCK 6  
 Db 298 dfgwck 303

RESULT 13

AAE00236

ID AAE00236 standard; Protein; 303 AA.

XX

AC AAE00236;

XX

DT 13-JUN-2001 (first entry)

XX

DE Protein encoded by Probe 10.

XX

KW Transacyclase; taxol; paclitaxel biosynthesis; taxoid; probe;

XX

KW transgenic organism; TAX10.

XX

OS Taxus cuspidata.

XX

PR WO200123586-A2.

XX

PN 05-APR-2001.

XX

PF 29-SEP-2000; 2000WO-US27006.

XX

PR 30-SEP-1999; 99US-0411145.

XX

PR 07-DEC-1999; 99US-0457046.

XX

PA (UNIW ) UNIV WASHINGTON STATE RES FOUND.

XX

PI Croteau RB, Walker KD, Schoendorf A, Wildung MR;

XX

PR N-PSDB; AAD03342.

XX

DR 05-APR-2001.

XX

PP 29-SEP-2000; 2000WO-US27006.

XX

PR 30-SEP-1999; 99US-0411145.

XX

PR 07-DEC-1999; 99US-0457046.

XX

PA (UNIW ) UNIV WASHINGTON STATE RES FOUND.

XX

PI Croteau RB, Walker KD, Schoendorf A, Wildung MR;

XX

PR N-PSDB; AAD03343.

XX

PT New transacyclase enzymes, useful for the high yield production of

PT Taxol(TM), related taxoids and useful intermediates in the in the

PT paclitaxel biosynthetic pathway -

XX

PS Claim 1; Page 96-97; 162pp; English.

XX

CC The present sequence is the amino acid sequence of probe 1.

CC Probe 1 is derived from AT-FOR1 and AT-REV1 primers and is used for screening

CC Taxus cuspidata full length cDNA clones TAX1 or TAX01 (taxadien-5 alpha-

CC ol acetyl transferase) and TAX2 or TAX02 (taxadien-2 alpha-O-benzoyl

CC transferase). The probes are useful for the identification of (nucleic

CC acid sequences encoding) transacyclases. The probes isolated from

CC the taxus genus are useful for the synthetic production of Taxol(TM) and

CC related taxoids, as well as intermediates in the paclitaxel biosynthetic

CC pathway. They can also be used for the creation of transgenic organisms

CC that either produce the transacyclases for subsequent in vitro use, or

CC produce the transacyclases in vivo. The (nucleic acids encoding)

CC transacyclases are also useful for creating specific binding agents that

CC recognise the corresponding transacyclases. Binding agents include

CC (fragments of) antibodies or any other agent capable of specifically

CC binding to the groups on the proteins.

Query Match 100.0%; Score 40; DB 22; Length 303;  
 Best Local Similarity 100.0%; Pred. No. 27; Mismatches 0; Indels 0; Gaps 0;  
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DFGWCK 6

Db 298 dfgwck 303

RESULT 14

AAE00227

ID AAE00227 standard; Protein; 306 AA.

XX

AC AAE00227;

XX

DT 13-JUN-2001 (first entry)

XX

DE Protein encoded by Probe 1.

XX

KW TAX1; taxadien-5 alpha-O-benzoyl transferase; paclitaxel biosynthesis;

XX

KW taxoid; transgenic organism; probe.

XX

OS Taxus cuspidata.

XX

PN WO200123586-A2.

XX

PD 05-APR-2001.

XX

PF 29-SEP-2000; 2000WO-US27006.

XX

PR 30-SEP-1999; 99US-0411145.

XX

PR 07-DEC-1999; 99US-0457046.

XX

PA (UNIW ) UNIV WASHINGTON STATE RES FOUND.

XX

PI Croteau RB, Walker KD, Schoendorf A, Wildung MR;

XX

PR N-PSDB; AAD03343.

XX

PT New transacyclase enzymes, useful for the high yield production of

PT Taxol(TM), related taxoids and useful intermediates in the in the

PT paclitaxel biosynthetic pathway -

XX

PS Claim 1; Page 96-97; 162pp; English.

XX

CC The present sequence is the amino acid sequence of probe 1.

CC Probe 1 is derived from AT-FOR1 and AT-REV1 primers and is used for screening

CC Taxus cuspidata full length cDNA clones TAX1 or TAX01 (taxadien-5 alpha-

CC ol acetyl transferase) and TAX2 or TAX02 (taxadien-2 alpha-O-benzoyl

CC transferase). The probes are useful for the identification of (nucleic

CC acid sequences encoding) transacyclases. The probes isolated from

CC the taxus genus are useful for the synthetic production of Taxol(TM) and

CC related taxoids, as well as intermediates in the paclitaxel biosynthetic

CC pathway. They can also be used for the creation of transgenic organisms

CC that either produce the transacyclases for subsequent in vitro use, or

CC produce the transacyclases in vivo. The (nucleic acids encoding)

CC transacyclases are also useful for creating specific binding agents that

CC recognise the corresponding transacyclases. Binding agents include

CC (fragments of) antibodies or any other agent capable of specifically

CC binding to the groups on the proteins.

XX

SQ Sequence 306 AA;

Query Match 100.0%; Score 40; DB 22; Length 306;

Best Local Similarity 100.0%; Pred. No. 27; Mismatches 0; Indels 0; Gaps 0;

Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DFGWCK 6

Db 298 dfgwck 303

SQ Sequence 303 AA:

Query Match 100.0%; Score 40; DB 22; Length 303;

Best Local Similarity 100.0%; Pred. No. 27; Mismatches 0; Indels 0; Gaps 0;

Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DFGWCK 6

Db 298 dfgwck 303

SQ Sequence 303 AA:

Query Match 100.0%; Score 40; DB 22; Length 303;

Best Local Similarity 100.0%; Pred. No. 27; Mismatches 0; Indels 0; Gaps 0;

Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DFGWCK 6

Db 298 dfgwck 303

SQ Sequence 303 AA:

Query Match 100.0%; Score 40; DB 22; Length 303;

Best Local Similarity 100.0%; Pred. No. 27; Mismatches 0; Indels 0; Gaps 0;

Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DFGWCK 6

Db 298 dfgwck 303

SQ Sequence 303 AA:

Query Match 100.0%; Score 40; DB 22; Length 303;

Best Local Similarity 100.0%; Pred. No. 27; Mismatches 0; Indels 0; Gaps 0;

Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DFGWCK 6

Db 298 dfgwck 303

SQ Sequence 303 AA:

Query Match 100.0%; Score 40; DB 22; Length 303;

Best Local Similarity 100.0%; Pred. No. 27; Mismatches 0; Indels 0; Gaps 0;

Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DFGWCK 6

Db 298 dfgwck 303

SQ Sequence 303 AA:

Query Match 100.0%; Score 40; DB 22; Length 303;

Best Local Similarity 100.0%; Pred. No. 27; Mismatches 0; Indels 0; Gaps 0;

Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DFGWCK 6

Db 298 dfgwck 303

SQ Sequence 303 AA:

Query Match 100.0%; Score 40; DB 22; Length 303;

Best Local Similarity 100.0%; Pred. No. 27; Mismatches 0; Indels 0; Gaps 0;

Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DFGWCK 6

Db 298 dfgwck 303

SQ Sequence 303 AA:

Query Match 100.0%; Score 40; DB 22; Length 303;

Best Local Similarity 100.0%; Pred. No. 27; Mismatches 0; Indels 0; Gaps 0;

Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DFGWCK 6

Db 298 dfgwck 303

SQ Sequence 303 AA:

Query Match 100.0%; Score 40; DB 22; Length 303;

Best Local Similarity 100.0%; Pred. No. 27; Mismatches 0; Indels 0; Gaps 0;

Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DFGWCK 6

Db 298 dfgwck 303

SQ Sequence 303 AA:

Query Match 100.0%; Score 40; DB 22; Length 303;

Best Local Similarity 100.0%; Pred. No. 27; Mismatches 0; Indels 0; Gaps 0;

Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DFGWCK 6

Db 298 dfgwck 303

SQ Sequence 303 AA:

Query Match 100.0%; Score 40; DB 22; Length 303;

Best Local Similarity 100.0%; Pred. No. 27; Mismatches 0; Indels 0; Gaps 0;

Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DFGWCK 6

Db 298 dfgwck 303

SQ Sequence 303 AA:

Query Match 100.0%; Score 40; DB 22; Length 303;

Best Local Similarity 100.0%; Pred. No. 27; Mismatches 0; Indels 0; Gaps 0;

Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DFGWCK 6

Db 298 dfgwck 303

SQ Sequence 303 AA:

Query Match 100.0%; Score 40; DB 22; Length 303;

Best Local Similarity 100.0%; Pred. No. 27; Mismatches 0; Indels 0; Gaps 0;

Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DFGWCK 6

Db 298 dfgwck 303

SQ Sequence 303 AA:

Query Match 100.0%; Score 40; DB 22; Length 303;

Best Local Similarity 100.0%; Pred. No. 27; Mismatches 0; Indels 0; Gaps 0;

Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DFGWCK 6

Db 298 dfgwck 303

SQ Sequence 303 AA:

Query Match 100.0%; Score 40; DB 22; Length 303;

Best Local Similarity 100.0%; Pred. No. 27; Mismatches 0; Indels 0; Gaps 0;

Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DFGWCK 6

Db 298 dfgwck 303

SQ Sequence 303 AA:

Query Match 100.0%; Score 40; DB 22; Length 303;

Best Local Similarity 100.0%; Pred. No. 27; Mismatches 0; Indels 0; Gaps 0;

Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DFGWCK 6

Db 298 dfgwck 303

SQ Sequence 303 AA:

Query Match 100.0%; Score 40; DB 22; Length 303;

Best Local Similarity 100.0%; Pred. No. 27; Mismatches 0; Indels 0; Gaps 0;

Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DFGWCK 6

Db 298 dfgwck 303

SQ Sequence 303 AA:

Query Match 100.0%; Score 40; DB 22; Length 303;

Best Local Similarity 100.0%; Pred. No. 27; Mismatches 0; Indels 0; Gaps 0;

Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DFGWCK 6

Db 298 dfgwck 303

SQ Sequence 303 AA:

Query Match 100.0%; Score 40; DB 22; Length 303;

Best Local Similarity 100.0%; Pred. No. 27; Mismatches 0; Indels 0; Gaps 0;

Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DFGWCK 6

Db 298 dfgwck 303

SQ Sequence 303 AA:

Query Match 100.0%; Score 40; DB 22; Length 303;

Best Local Similarity 100.0%; Pred. No. 27; Mismatches 0; Indels 0; Gaps 0;

Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DFGWCK 6

Db 298 dfgwck 303

SQ Sequence 303 AA:

Query Match 100.0%; Score

Db 301 dfgwgk 306

RESULT 15  
 AAE00228  
 ID AAE00228 standard; Protein; 306 AA.  
 XX  
 AC AAE00228;  
 XX  
 DT 13-JUN-2001 (first entry)  
 DE Protein encoded by Probe 2.  
 XX  
 KW TAX2; taxadien-2 alpha-O-benzoyl transferase; paclitaxel biosynthesis;  
 KW taxoid; transgenic organism; probe; transacylase; taxol.  
 XX  
 OS Taxus cuspidata.  
 XX  
 PN WO200123586-A2.  
 XX  
 PD 05-APR-2001.  
 XX  
 PR 29-SEP-2000; 20000R0-US27006.  
 XX  
 PR 30-SEP-1999; 99US-0411145.  
 PR 07-DEC-1999; 99US-0457046.  
 XX  
 PA (UNIW ) UNIV WASHINGTON STATE RES FOUND.  
 XX  
 PT Croteau RB, Walker KD, Schoendorf A, Wildung MR;  
 XX  
 DR WPI; 2001-245004/25.  
 DR N-PSDB; AAD0334.

PT New transacylase enzymes, useful for the high yield production of  
 PT TAXOL™, related taxoids and useful intermediates in the in the  
 PT paclitaxel biosynthetic pathway

XX  
 PS Claim 1; Page 98-99; 162pp; English.

CC The present sequence is the amino acid sequence of probe 2. Probe 2  
 CC is used for screening Taxus cuspidata transacylase cDNA. Probe 2  
 CC is derived from AT-FOR1 and AT-REV1 primers but it is not useful  
 CC for screening TAX2 or TAX02 (taxadien-2 alpha-O-benzoyl  
 CC transferase) but has the sequence similar to TAX2 gene. The  
 CC probes are useful for the identification of (nucleic acid sequences  
 CC encoding) transacylases. The probes isolated from the Taxus genus  
 CC are useful for the synthetic production of TAXOL™ and related taxoids,  
 CC as well as intermediates in the paclitaxel biosynthetic pathway. They  
 CC can also be used for the creation of transgenic organisms that either  
 CC produce the transacylases for subsequent *in vitro* use, or produce the  
 CC transacylases *in vivo*. The (nucleic acids encoding) transacylases are  
 CC also useful for creating specific binding agents that recognise the  
 CC corresponding transacylases. Binding agents include (fragments of)  
 CC antibodies or any other agent capable of specifically binding to the  
 CC groups on the proteins.  
 XX  
 Sequence 306 AA;

SQ

Query Match 100.0%; Score 40; DB 22; Length 306;  
 Best Local Similarity 100.0%; Pred. No. 27;  
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY	1 DFGWGK 6
Db	301 dfgwgk 306

Search completed: November 5, 2001, 12:57:31  
 Job time: 33 sec



Query Match 100.0%; Score 40; DB 4; Length 302;  
 Best Local Similarity 100.0%; Pred. No. 9.3; CURRENT APPLICATION NUMBER: US/09/457, 046B  
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 SEQ ID NO: 18  
 LENGTH: 302

RESULT 3  
 US-09-457-046B-10  
 ; Sequence 10, Application US/09457046B  
 ;  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Croteau, Rodney et al.  
 ; TITLE OF INVENTION: Transacylases of the Paclitaxel Biosynthetic Pathway  
 ; FILE REFERENCE: 53679  
 ; CURRENT APPLICATION NUMBER: US/09/457, 046B  
 ; CURRENT FILING DATE: 1999-12-07  
 ; NUMBER OF SEQ ID NOS: 74  
 ; SOFTWARE: Patentin Ver. 2.1  
 ; SEQ ID NO: 10  
 ; LENGTH: 302  
 ; TYPE: PRT  
 ; ORGANISM: Taxus cuspidata  
 ; US-09-457-046B-10

Query Match 100.0%; Score 40; DB 4; Length 302;  
 Best Local Similarity 100.0%; Pred. No. 9.3; CURRENT APPLICATION NUMBER: US/09/457, 046B  
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 SEQ ID NO: 18  
 LENGTH: 302

RESULT 4  
 US-09-457-046B-16  
 ; Sequence 16, Application US/09457046B  
 ;  
 ; PATENT NO: 6287835  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Croteau, Rodney et al.  
 ; TITLE OF INVENTION: Transacylases of the Paclitaxel Biosynthetic Pathway  
 ; FILE REFERENCE: 53679  
 ; CURRENT APPLICATION NUMBER: US/09/457, 046B  
 ; CURRENT FILING DATE: 1999-12-07  
 ; NUMBER OF SEQ ID NOS: 74  
 ; SOFTWARE: Patentin Ver. 2.1  
 ; SEQ ID NO: 16  
 ; LENGTH: 302  
 ; TYPE: PRT  
 ; ORGANISM: Taxus cuspidata  
 ; US-09-457-046B-24

Query Match 100.0%; Score 40; DB 4; Length 302;  
 Best Local Similarity 100.0%; Pred. No. 9.3; CURRENT APPLICATION NUMBER: US/09/457, 046B  
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 SEQ ID NO: 18  
 LENGTH: 302

RESULT 5  
 US-09-457-046B-16  
 ; Sequence 18, Application US/09457046B  
 ;  
 ; PATENT NO: 6287835  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Croteau, Rodney et al.  
 ; TITLE OF INVENTION: Transacylases of the Paclitaxel Biosynthetic Pathway  
 ; FILE REFERENCE: 53679  
 ; CURRENT APPLICATION NUMBER: US/09/457, 046B  
 ; CURRENT FILING DATE: 1999-12-07  
 ; NUMBER OF SEQ ID NOS: 74  
 ; SOFTWARE: Patentin Ver. 2.1  
 ; SEQ ID NO: 16  
 ; LENGTH: 302  
 ; TYPE: PRT  
 ; ORGANISM: Taxus cuspidata  
 ; US-09-457-046B-12

Query Match 100.0%; Score 40; DB 4; Length 302;  
 Best Local Similarity 100.0%; Pred. No. 9.3; CURRENT APPLICATION NUMBER: US/09/457, 046B  
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 SEQ ID NO: 18  
 LENGTH: 302

RESULT 6  
 US-09-457-046B-24  
 ; Sequence 24, Application US/09457046B  
 ;  
 ; PATENT NO: 6287835  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Croteau, Rodney et al.  
 ; TITLE OF INVENTION: Transacylases of the Paclitaxel Biosynthetic Pathway  
 ; FILE REFERENCE: 53679  
 ; CURRENT APPLICATION NUMBER: US/09/457, 046B  
 ; CURRENT FILING DATE: 1999-12-07  
 ; NUMBER OF SEQ ID NOS: 74  
 ; SOFTWARE: Patentin Ver. 2.1  
 ; SEQ ID NO: 24  
 ; LENGTH: 302  
 ; TYPE: PRT  
 ; ORGANISM: Taxus cuspidata  
 ; US-09-457-046B-24

Query Match 100.0%; Score 40; DB 4; Length 302;  
 Best Local Similarity 100.0%; Pred. No. 9.3; CURRENT APPLICATION NUMBER: US/09/457, 046B  
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 SEQ ID NO: 18  
 LENGTH: 302

RESULT 7  
 US-09-457-046B-12  
 ; Sequence 12, Application US/09457046B  
 ;  
 ; PATENT NO: 6287835  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Croteau, Rodney et al.  
 ; TITLE OF INVENTION: Transacylases of the Paclitaxel Biosynthetic Pathway  
 ; FILE REFERENCE: 53679  
 ; CURRENT APPLICATION NUMBER: US/09/457, 046B  
 ; CURRENT FILING DATE: 1999-12-07  
 ; NUMBER OF SEQ ID NOS: 74  
 ; SOFTWARE: Patentin Ver. 2.1  
 ; SEQ ID NO: 12  
 ; LENGTH: 303  
 ; TYPE: PRT  
 ; ORGANISM: Taxus cuspidata  
 ; US-09-457-046B-12

Query Match 100.0%; score 40; DB 4; Length 303;  
 Best Local Similarity 100.0%; pred. No. 9.3; LENGTH: 303

```

Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 1 DFGW GK 6
Db 298 DFGW GK 303
RESULT 8
US-09-457-046B-20
; Sequence 20, Application US/09457046B
; Patent No. 6287835
; GENERAL INFORMATION:
; APPLICANT: Croteau, Rodney et al.
; TITLE OF INVENTION: Transacyclases of the Paclitaxel Biosynthetic Pathway
; FILE REFERENCE: 53679
; CURRENT APPLICATION NUMBER: US/09/457,046B
; CURRENT FILING DATE: 1999-12-07
; NUMBER OF SEQ ID NOS: 74
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 20
; LENGTH: 303
; TYPE: PRT
; ORGANISM: Taxus cuspidata
; US-09-457-046B-20

RESULT 11
US-09-457-046B-22
; Sequence 22, Application US/09457046B
; Patent No. 6287835
; GENERAL INFORMATION:
; APPLICANT: Croteau, Rodney et al.
; TITLE OF INVENTION: Transacyclases of the Paclitaxel Biosynthetic Pathway
; FILE REFERENCE: 53679
; CURRENT APPLICATION NUMBER: US/09/457,046B
; CURRENT FILING DATE: 1999-12-07
; NUMBER OF SEQ ID NOS: 74
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 22
; LENGTH: 306
; TYPE: PRT
; ORGANISM: Taxus cuspidata
; US-09-457-046B-22

RESULT 9
US-09-457-046B-2
; Sequence 2, Application US/09457046B
; Patent No. 6287835
; GENERAL INFORMATION:
; APPLICANT: Croteau, Rodney et al.
; TITLE OF INVENTION: Transacyclases of the Paclitaxel Biosynthetic Pathway
; FILE REFERENCE: 53679
; CURRENT APPLICATION NUMBER: US/09/457,046B
; CURRENT FILING DATE: 1999-12-07
; NUMBER OF SEQ ID NOS: 74
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 2
; LENGTH: 306
; TYPE: PRT
; ORGANISM: Taxus cuspidata
; US-09-457-046B-2

RESULT 10
Query Match 100.0%; Score 40; DB 4; Length 306;
Best Local Similarity 100.0%; Pred. No. 9.4; Mismatches 0; Indels 0; Gaps 0;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 1 DFGW GK 6
Db 301 DFGW GK 306
Db 301 DFGW GK 306
RESULT 10
US-09-457-046B-4
; Sequence 4, Application US/09457046B
; Patent No. 6287835
; GENERAL INFORMATION:
; APPLICANT: Croteau, Rodney et al.
; TITLE OF INVENTION: Transacyclases of the Paclitaxel Biosynthetic Pathway
; FILE REFERENCE: 53679
; CURRENT APPLICATION NUMBER: US/09/457,046B
; CURRENT FILING DATE: 1999-12-07
; NUMBER OF SEQ ID NOS: 74
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 14
; LENGTH: 322
; TYPE: PRT
; ORGANISM: Taxus cuspidata
; US-09-457-046B-4

RESULT 12
Query Match 100.0%; Score 40; DB 4; Length 322;
Best Local Similarity 100.0%; Pred. No. 9.8; Mismatches 0; Indels 0; Gaps 0;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 1 DFGW GK 6
Db 317 DFGW GK 322
Db 317 DFGW GK 322

```

RESULT 13

US-09-457-046B-59

Sequence 59, Application US/09457046B

Patent No. 6287835

GENERAL INFORMATION:

APPLICANT: Croteau, Rodney et al.

FILE REFERENCE: 53679

CURRENT APPLICATION NUMBER: US/09/457,046B

CURRENT FILING DATE: 1999-12-07

NUMBER OF SEQ ID NOS: 74

SOFTWARE: PatentIn Ver. 2.1

SEQ ID NO: 59

LENGTH: 331

TYPE: PRT

ORGANISM: Arabidopsis thaliana

US-09-457-046B-59

Query Match 100.0%; Score 40; DB 4; Length 439;

Best Local Similarity 100.0%; Pred. No. 13; Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 DFGWCK 6  
Db 265 DFGWCK 270

RESULT 14

US-09-457-046B-62

Sequence 62, Application US/09457046B

Patent No. 6287835

GENERAL INFORMATION:

APPLICANT: Croteau, Rodney et al.

TITLE OF INVENTION: Transacylases of the Paclitaxel Biosynthetic Pathway

FILE REFERENCE: 53679

CURRENT APPLICATION NUMBER: US/09/457,046B

CURRENT FILING DATE: 1999-12-07

NUMBER OF SEQ ID NOS: 74

SOFTWARE: PatentIn Ver. 2.1

SEQ ID NO: 62

LENGTH: 436

TYPE: PRT

ORGANISM: Arabidopsis thaliana

US-09-457-046B-62

Query Match 100.0%; Score 40; DB 4; Length 439;

Best Local Similarity 100.0%; Pred. No. 13; Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 DFGWCK 6  
Db 372 DFGWCK 377

RESULT 15

US-09-457-046B-68

Sequence 68, Application US/09457046B

Patent No. 6287835

GENERAL INFORMATION:

APPLICANT: Croteau, Rodney et al.

TITLE OF INVENTION: Transacylases of the Paclitaxel Biosynthetic Pathway

FILE REFERENCE: 53679

CURRENT APPLICATION NUMBER: US/09/457,046B

CURRENT FILING DATE: 1999-12-07

NUMBER OF SEQ ID NOS: 74

SOFTWARE: PatentIn Ver. 2.1

SEQ ID NO: 68

LENGTH: 439

TYPE: PRT

ORGANISM: Catharanthus roseus

US-09-457-046B-68

Query Match 100.0%; Score 40; DB 4; Length 439;

Best Local Similarity 100.0%; Pred. No. 13; Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 DFGWCK 6  
Db 380 DFGWCK 385Search completed: November 5, 2001, 12:58:09  
Job time: 70 sec

Mon Dec 10 13:33:28 2001

us-08-894-356c-21.rai

**THIS PAGE BLANK (USPTO)**

GenCore - version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.  
OM protein - protein search, using sw model

Run on:	November 5, 2001, 12:56:59 ;	Search time 12.79 Seconds (without alignments)
		35.1735 Million cell updates/sec
Title:	US-08-894-356C-21	
Perfect score:	40	
Sequence:	1 DFGWGR 6	
Scoring table:	BLOSUM62	
	Capop 10.0 , Gapext 0.5	
Searched:	219241 seqs, 76174552 residues	
Total number of hits satisfying chosen parameters:	219241	
Minimum DB seq length:	0	
Maximum DB seq length:	200000000	
Post-processing:	Maximum Match 0% Listing first 45 summaries	
Database :	PIR_68:*	
	1: pirl:*	
	2: pir2:*	
	3: pir3:*	
	4: pir4:*	
Pred. No.	No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.	
SUMMARIES		
Result No.	Score	Query Match Length DB ID Description
1	40	100 0 436 2 D86378 Protein F21J9.8 [1
2	40	100 0 443 2 T45574 anthranilate N-hyd
3	40	100 0 450 2 T45573 anthranilate N-hyd
4	40	100 0 451 2 T00527 hypothetical prote
5	40	100 0 455 2 T09666 probable anthranil
6	40	100 0 460 2 T03274 hsr201 protein, hy
7	40	100 0 461 2 T00918 hypothetical prote
8	40	100 0 475 2 T45961 anthranilate N-hyd
9	37	92 5 399 2 F83484 probable MFS trans
10	37	92 5 482 2 G84823 probable anthocyan
11	36	90 0 430 2 T46216 hypothetical prote
12	35	87 5 439 2 T52321 taxadienol acetyl
13	35	87 5 440 2 T52320 10-deacetylbaacci
14	35	87 5 442 2 T10718 anthranilate N-ben
15	35	87 5 445 2 T10717 anthranilate N-ben
16	35	87 5 445 2 T10719 anthranilate N-ben
17	35	87 5 446 2 T10711 anthranilate N-ben
18	35	87 5 464 2 G86453 YUP8H12R_39, homol
19	35	87 5 476 2 D84505 probable membrane
20	35	87 5 480 2 R86411 protein FK23_12_1
21	35	87 5 485 2 T33865 hypothetical prote
22	35	87 5 572 2 T013865 hypothetical prote
23	34	85 0 113 2 S30515 submitted to the Protein Sequence Database, December 1999
24	34	85 0 129 2 S60978 wound-induced prot
25	34	85 0 215 2 C85327 probable membrane
26	34	85 0 216 2 T02913 hypothetical prote
27	34	85 0 303 2 B83336 hypothetical prote
28	34	85 0 493 2 G84594 probable diacylglycerol
29	34	85 0 519 2 S54300 transketolase (EC

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

RESULT	1	
D86378	Protein F21J9.8 [1	C;Species: Arabidopsis thaliana
	protein F21J9.8 [imported] - Arabidopsis thaliana (mouse/ear cress)	C;Accession: D86378
	R;Theologis, A.; Ecker, J.R.; Palm, C.J.; Federer, N.A.; Kaul, S.; White, O.; Alon	C;Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 31-Mar-2001
	Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar,	
	ansen, N.F.; Hughes, B.; Huijzer, L.	
	Nature 408, 816-820, 2000	
	A;Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Luos, J.S.; Maiti, R.; Marzia	
	Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.	
	A;Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallo	
	ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.	
	A;Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.	
	A;Reference number: A86141; PMID: 11061719	
	A;Accession: D86378	
	A;Status: preliminary	
	A;Molecule type: DNA	
	A;Residues: 1-436 <STO>	
	A;Cross-references: GB:AB005172; NTID:9743331; PIDN:AAF97955.1; GSPDB:GN00141	
	C;Genetics:	
	A;Gene: F21J9.8	
	A;Map position: 1	
QY	1 DFGWGR 6	
Db	372 DFGWGR 377	
RESULT		
T45574	anthranilate N-hydroxycinnamoyl/benzoyltransferase-like protein - Arabidopsis thaliana	C;Species: Arabidopsis thaliana (mouse/ear cress)
	N;Alternate names: protein F11C1_120	C;Accession: T45574
	K;Burgess, M.; Collado, M.C.; Navarro, P.; Terol, J.; Perez-Alonso, M.; Mewes, H.W.; A;Reference number: 223007	C;Date: 04-Feb-2000 #sequence_revision 04-Feb-2000 #text_change 04-Feb-2000
	A;Accession: T45574	
	A;Status: preliminary	
	A;Molecule type: DNA	
	A;Residues: 1-443 <BAR>	
	A;Cross-references: EMBL:AL132976	
	A;Experimental source: cultivar Columbia; BAC clone F11C1	

C;Genetics:  
A;Map position: 3  
A;Map: 3  
A;Note: F1C1.1120

Query Match ;  
Best Local Similarity 100.0%; Score 40; DB 2; Length 443;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 DFGWCK 6  
Db 386 DFGWCK 391

RESULT 3

T45533 anthranilate N-hydroxycinnamoyl/benzoyltransferase-like protein - Arabidopsis thaliana  
N;Alternate names: protein F1C1.110  
C;Species: Arabidopsis thaliana (mouse-ear cress)  
C;Accession: T45533 04-Feb-2000 #sequence\_revision 04-Feb-2000  
R;Baugues, M.; Collado, M.C.; Navarro, P.; Terol, J.; Perez-Alonso, M.; Mewes, H.W.; May submitted to the Protein Sequence Database, December 1999  
A;Reference number: Z23007  
A;Accession: T45573  
A;Status: preliminary  
A;Molecule type: DNA  
A;Residues: 1-450 <BAR>  
A;Cross-references: EMBL:AL132976  
A;Experimental source: cultivar Columbia; BAC clone F1C1  
C;Genetics:  
A;Map position: 3  
A;Note: F1C1.110

Query Match ;  
Best Local Similarity 100.0%; Score 40; DB 2; Length 450;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 DFGWCK 6  
Db 391 DFGWCK 396

RESULT 4

T00271 hypothetical protein At2g19070 [imported] - Arabidopsis thaliana  
N;Alternate names: hypothetical protein T20k24.8  
C;Species: Arabidopsis thaliana (mouse-ear cress)  
C;Accession: T00271; B84572  
R;Rounsley, S.D.; Kaul, S.; Lin, X.; Ketchum, K.A.; Crosby, M.L.; Brandon, R.C.; Sykes, submitted to the EMBL Data Library; July 1997  
A;Description: Arabidopsis thaliana chromosome II BAC T20k24 genomic sequence.  
A;Reference number: Z14167  
A;Accession: T00527  
A;Status: translated from GB/EMBL/DDBJ  
A;Molecule type: DNA  
A;Residues: 1-451 <ROU>  
A;Cross-references: EMBL:AC002392; NID:g3176701; PID:93176709  
A;Experimental source: Cultivar Columbia  
R;Lin, X.; Kaul, S.; Rounsley, S.D.; Shea, T.P.; Benito, M.I.; Town, C.D.; Fujii, C.Y.; M.; Kao, H.; Moffat, K.S.; Cronin, L.A.; Shen, M.; Vanken, S.E.; Umayam, L.; Tallon, L.; euss, D.; Nieman, W.C.; White, O.; Elsen, J.A.; Salzberg, S.L.; Fraser, C.M.; Venter, J. Nature 402, 765-768, 1999  
A;Title: Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana.  
A;Reference number: A84420; MID:20083487  
A;Accession: B84572  
A;Status: preliminary  
A;Molecule type: DNA  
A;Residues: 1-451 <STO>  
A;Cross-references: GB:AP002093; NID:g3176709; PIDN:AAD12025.1; GSPDB:GN00139  
C;Genetics:

A;Gene: At2g19070; T20K24.8  
A;Map Position: 2  
A;Introns: 322/1

Query Match ;  
Best Local Similarity 100.0%; Score 40; DB 2; Length 451;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 DFGWCK 6  
Db 397 DFGWCK 402

RESULT 5

T09666 anthranilate N-benzoyltransferase (EC 2.3.1.144) - muskmelon (fragment)  
N;Alternate names: anthranilate N-hydroxycinnamoyl/benzoyltransferase  
C;Species: Cucumis melo (musk-melon)  
C;Accession: T09666 #sequence\_revision 16-Jul-1999 #text\_change 21-Jul-2000  
R;Agellis, A.; John, I.; Karvouni, Z.; Grierson, D. Plant Mol. Biol. 33, 313-322, 1997  
A;Title: Characterization of two cDNA clones for mRNAs expressed during ripening of muskmelons  
A;Reference number: Z16810; MID:97188564  
A;Accession: T09666  
A;Status: Preliminary; translated from GB/EMBL/DDBJ  
A;Molecule type: mRNA  
A;Residues: 1-455 <AGG>  
A;Cross-references: EMBL:270521; NID:g1843439; PIDN:CAA94432.1; PID:g1843440  
A;Experimental source: cultivar Cantaloupe charentais; pericarp of ripe fruit  
C;Keywords: acyltransferase; coenzyme A

Query Match ;  
Best Local Similarity 100.0%; Score 40; DB 2; Length 455;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 DFGWCK 6  
Db 374 DFGWCK 379

RESULT 6

T03274 hsr201 protein, hypersensitivity-related - common tobacco  
C;Species: Nicotiana tabacum (common tobacco)  
C;Accession: T03274 #sequence\_revision 24-Mar-1999 #text\_change 21-Jul-2000  
R;Cernic, P.; Huang, H.-C.; Marco, Y. Plant Mol. Biol. 31, 255-265, 1996  
A;Title: Characterization of hsr201 and hsr515, two tobacco genes preferentially expressed  
A;Reference number: Z14876; MID:96343929  
A;Accession: T03274  
A;Status: Preliminary; translated from GB/EMBL/DDBJ  
A;Molecule type: mRNA  
A;Residues: 1-460 <CGZ>  
A;Cross-references: EMBL:X95343; NID:g11171576; PIDN:CAA64636.1; PID:g11171577  
A;Experimental source: cultivar bottom special  
C;Genetics:  
A;Gene: hsr201

Query Match ;  
Best Local Similarity 100.0%; Score 40; DB 2; Length 460;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 DFGWCK 6  
Db 382 DFGWCK 387

RESULT 7

T00918 hypothetical protein F21B7\_32 - Arabidopsis thaliana  
 C;Species: Arabidopsis thaliana (mouse-ear cress)  
 C;Date: 12-Feb-1999 #sequence\_revision 12-Feb-1999 #text\_change 22-Oct-1999  
 C;Accession: T00918  
 R;Shinn, P.; Buehler, E.; Dewar, K.; Feng, J.; Kim, C.; Li, Y.; Sun, H.; Conway, A.; Cologe, A.; Ecker, J.R.  
 submitted to the EMBL Data Library, January 1998  
 A;Description: Genomic sequence for Arabidopsis thaliana BAC F21B7.  
 A;Reference number: 214208  
 A;Status: translated from GB/EMBL/DDJB  
 A;Molecule type: DNA  
 A;Residues: 1-461 <SHT>  
 A;Cross-references: EMBL:AC002560; NID:92618677; PID:92809263; GSPDB:GN00059; ATSP:F21B7  
 C;Genetics:  
 A;Gene: ATSP:F21B7\_32  
 A;Map position: 1

RESULT 8

Query Match 100.0%; Score 40; DB 2; Length 461;  
 Best Local Similarity 100.0%; Pred. No. 11; Mismatches 0; Indels 0; Gaps 0;  
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 DFGW GK 6  
 Db 400 DFGW GK 405

Query Match 92.5%; Score 37; DB 2; Length 399;  
 Best Local Similarity 83.3%; Pred. No. 32; Mismatches 0; Indels 0; Gaps 0;  
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 DFGW GK 6  
 Db 36 DFGW GR 41

RESULT 9

T45961

anthranilate N-benzoyltransferase-like protein - Arabidopsis thaliana  
 N;Alternate name: protein F7J8\_190  
 C;Species: Arabidopsis thaliana (mouse-ear cress)  
 C;Date: 04-Feb-2000 #sequence\_revision 04-Feb-2000 #text\_change 04-Feb-2000  
 C;Accession: T45961  
 R;Bevan, M.; Zimmermann, W.; Grueneisen, A.; Wambutt, R.; Bancroft, I.; Meves, H.W.; Lern  
 submitted to the Protein Sequence Database, January 2000  
 A;Reference number: 223018  
 A;Accession: 143961  
 A;Status: preliminary  
 A;Molecule type: DNA  
 A;Residues: 1-475 <BEV>  
 A;Cross-references: EMBL:AL137189  
 A;Experimental source: cultivar Columbia; BAC clone F7J8  
 C;Genetics:  
 A;Map position: 5  
 A;Note: F7J8\_190

T45961

Query Match 100.0%; Score 40; DB 2; Length 475;  
 Best Local Similarity 100.0%; Pred. No. 11; Mismatches 0; Indels 0; Gaps 0;  
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 DFGW GK 6  
 Db 411 DFGW GK 416

RESULT 9

F83484

probable MFS transporter PA1286 [imported] - Pseudomonas aeruginosa (strain PA01)  
 C;Species: Pseudomonas aeruginosa  
 C;Date: 15-Sep-2000 #sequence\_revision 15-Sep-2000 #text\_change 31-Dec-2000  
 C;Accession: F83484  
 R;Stover, C.K.; Pham, X.Q.; Erwin, A.L.; Mizoguchi, S.D.; Warrener, P.; Hickey, M.J.; Br  
 adamson, S.; Yuan, Y.; Brody, L.L.; Coulter, S.N.; Folger, K.R.; Kas, A.; Larkby, K.; Lim,  
 ; Lory, S.; Olson, M.V.  
 Nature 406, 959-964, 2000  
 A;Title: Complete genome sequence of Pseudomonas aeruginosa PA01, an opportunistic patho  
 A;Reference number: A82950; MUID:2043737  
 A;Accession: F83484  
 A;Status: preliminary

A;Molecule type: DNA  
 A;Residues: 1-399 <STOP>  
 A;Cross-references: GB:AE00455B; GB:AE004091; NID:9947217; PIDN:AAG04675.1; GSPDB:GN  
 A;Experimental source: strain PA01  
 C;Genetics:  
 A;Gene: PA1286  

RESULT 10

G34823

probable anthocyanin 5-aromatic acyltransferase [imported] - Arabidopsis thaliana  
 C;Species: Arabidopsis thaliana (mouse-ear cress)  
 C;Date: 02-Feb-2001 #sequence\_revision 02-Feb-2001 #text\_change 02-Feb-2001  
 C;Accession: G34823  
 R;Lin, X.; Kaul, S.; Rounseley, S.D.; Shea, T.P.; Benito, M.I.; Town, C.D.; Fujii, C.Y  
 M.; Koo, H.; Moffat, K.S.; Cronin, L.A.; Shen, M.; VanAken, S.E.; Umavam, L.; Tallon,  
 euss, D.; Neerman, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Fraser, C.M.; Venter  
 Nature 402, 761-768, 1999  
 A;Title: Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana  
 A;Reference number: A84420; MUID:20083487  
 A;Accession: G34823  
 A;Status: preliminary  
 A;Molecule type: DNA  
 A;Residues: 1-462 <STOP>  
 A;Cross-references: GB:AE002093; NID:9208651; PIDN:AAB95283.1; GSPDB:GN00139  
 A;Gene: AT2g39980  
 A;Map position: 2

T46216

hypothetical protein T8P19\_230 - Arabidopsis thaliana  
 C;Species: Arabidopsis thaliana (mouse-ear cress)  
 C;Date: 04-Feb-2000 #sequence\_revision 04-Feb-2000 #text\_change 17-Mar-2000  
 C;Accession: T46216  
 R;Choisne, N.; Robert, C.; Brottier, P.; Wincker, P.; Cattolico, L.; Artiguenave, F.  
 A;Reference number: 223008  
 A;Accession: T46216  
 A;Status: preliminary  
 A;Molecule type: DNA  
 A;Residues: 1-430 <CHO>  
 A;Cross-references: EMBL:AL133315  
 A;Experimental source: cultivar Columbia; BAC clone T8P19  
 C;Genetics:  
 A;Map position: 3  
 A;Introns: 144/3  
 A;Note: T8P19\_230  
 C;Superfamily: Arabidopsis CER2 protein

T8P19\_230

Query Match 90.0%; Score 36; DB 2; Length 430;  
 Best Local Similarity 83.3%; Pred. No. 51; Mismatches 0; Indels 0; Gaps 0;

Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 DFGWCK 6  
 ||||:  
 Db 378 DFGWCE 383

RESULT 12  
 T52221 taxadienol acetyl transferase [imported] - *Taxus cuspidata*  
 C;Species: *Taxus cuspidata*  
 C;Date: 20-Oct-2000 #sequence\_revision 20-Oct-2000 #text\_change 20-Oct-2000  
 C;Accession: T52321  
 R;Walker, K.; Schoendorf, A.; Croteau, R.  
 submitted to the EMBL Data Library; September 1999  
 A;Description: Molecular Cloning of a Taxa<sub>4</sub>(20),11(12)-dien-5(alpha)-ol-O-Acetyl trans-  
 A;Reference number: Z260346  
 A;Accession: T52321  
 A;Status: preliminary; translated from GB/EMBL/DDBJ  
 A;Molecule type: mRNA  
 A;Residues: 1-439 <WAL>  
 A;Cross-references: EMBL:AF190130; PIDN:AAF34254.1  
 C;Genetics:  
 A;Gene: TAT

Query Match 87.5%; Score 35; DB 2; Length 439;  
 Best Local Similarity 100.0%; Pred. No. 77; Mismatches 5; Conservative 0; Indels 0; Gaps 0;  
 Matches 5; Mismatches 0;

Qy 1 DFGWCG 5  
 ||||:  
 Db 373 DFGWCG 377

RESULT 13  
 T52220 10-deacetylbaaccatin III-10-O-acetyl transferase [imported] - *Taxus cuspidata*  
 C;Species: *Taxus cuspidata*  
 C;Date: 20-Oct-2000 #sequence\_revision 20-Oct-2000 #text\_change 20-Oct-2000  
 C;Accession: T52320  
 R;Walker, K.; Croteau, R.  
 Proc. Natl. Acad. Sci. U.S.A. 97, 583-587, 2000  
 A;Title: Molecular cloning of a 10-deacetylbaaccatin III-10-O-acetyl transferase cDNA from  
 A;Reference number: Z260345  
 A;Accession: T52320  
 A;Status: preliminary; translated from GB/EMBL/DDBJ  
 A;Molecule type: mRNA  
 A;Residues: 1-440 <WAL>  
 A;Cross-references: EMBL:AF193765; PIDN:AAF27621.1  
 C;Genetics:  
 A;Gene: DBAT

Query Match 87.5%; Score 35; DB 2; Length 440;  
 Best Local Similarity 100.0%; Pred. No. 77; Mismatches 5; Conservative 0; Indels 0; Gaps 0;  
 Matches 5; Mismatches 0;

Qy 1 DFGWCG 5  
 ||||:  
 Db 372 DFGWCG 376

RESULT 14  
 T10718 antianilinate N-benzoyltransferase (EC 2.3.1.144) (clone pchcbtl) - clove pink  
 N;Alternate names: antranilate N-hydroxycinnamoyl/benzoyltransferase  
 C;Species: *Dianthus carophyllus* (clove pink)  
 C;Date: 16-Jul-1999 #sequence\_revision 16-Jul-1999 #text\_change 21-Jul-2000  
 C;Accession: T10718  
 R;Yang, Q.; Reinhard, K.; Schiltz, E.; Matern, U.  
 Plant Mol. Biol. 35, 777-789, 1997  
 A;Title: Characterization and heterologous expression of hydroxycinnamoyl/benzoyl-CoA transferase  
 A;Reference number: Z17095; MUID:98088004  
 A;Accession: T10717  
 A;Status: preliminary; translated from GB/EMBL/DDBJ  
 A;Molecule type: mRNA  
 A;Residues: 1-445 <WAN>  
 A;Cross-references: EMBL:Z84383; PIDN:CA06427.1; PID:92239083  
 C;Function:  
 C;Description: catalyzes the synthesis of antranilate  
 A;Pathway: phytalexin biosynthesis  
 C;Keywords: acyltransferase; coenzyme A

Query Match 87.5%; Score 35; DB 2; Length 445;  
 Best Local Similarity 100.0%; Pred. No. 78; Mismatches 5; Conservative 0; Indels 0; Gaps 0;  
 Matches 5; Mismatches 0;

Qy 1 DFGWCG 5  
 ||||:  
 Db 392 DFGWCG 396

Search completed: November 5, 2001, 12:57:51  
 Job time: 52 sec

RESUL

• • •

Query Match 87.5%; Score 35; DB 2; Length 440;  
 Best Local Similarity 100.0%; Pred. No. 77; Mismatches 5; Conservative 0; Indels 0; Gaps 0;  
 Matches 5; Mismatches 0;

Qy 1 DFGWCG 5  
 ||||:  
 Db 372 DFGWCG 376

RESULT 14  
 T10718 antianilinate N-benzoyltransferase (EC 2.3.1.144) (clone pchcbtl) - clove pink (fragment)  
 N;Alternate names: antranilate N-hydroxycinnamoyl/benzoyltransferase  
 C;Species: *Dianthus carophyllus* (clove pink)  
 C;Date: 16-Jul-1999 #sequence\_revision 16-Jul-1999 #text\_change 21-Jul-2000  
 C;Accession: T10718  
 R;Yang, Q.; Reinhard, K.; Schiltz, E.; Matern, U.  
 Plant Mol. Biol. 35, 777-789, 1997  
 A;Title: Characterization and heterologous expression of hydroxycinnamoyl/benzoyl-CoA transferase  
 A;Reference number: Z17095; MUID:98088004  
 A;Accession: T10717  
 A;Status: preliminary; translated from GB/EMBL/DDBJ  
 A;Molecule type: mRNA  
 A;Residues: 1-445 <WAN>  
 A;Cross-references: EMBL:Z84383; PIDN:CA06427.1; PID:92239083  
 C;Function:  
 C;Description: catalyzes the synthesis of antranilate  
 A;Pathway: phytalexin biosynthesis  
 C;Keywords: acyltransferase; coenzyme A

Mon Dec 10 13:33:29 2001

us-08-894-356c-21.rpr

**HIS PAGE BLANK (USPTO)**

GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

## OM protein - protein search, using sw model

Run on: November 5, 2001, 12:56:59 ; Search. time 10.15 seconds

(without alignments)  
21.674 Million cell updates/secTitle: US-08-894-356c-21  
Sequence: 40 DFGWKGScoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 100059 seqs, 36664827 residues

Total number of hits satisfying chosen parameters: 100059

Minimum DB seq length: 0  
Maximum DB seq length: 200000000Post-processing: Minimum Match 0%  
Maximum Match 100%

Listing first 45 summaries

Database : SwissProt\_39:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match Length	DB ID	Description
1	35	87.5	918	HXKL_BOVIN
2	35	87.5	1045	Y052_HUMAN
3	34	85.0	129	Y009_YEAST
4	34	85.0	361	G921_HUMAN
5	34	85.0	519	TKTC_CRAPL
6	34	85.0	667	TK71_BACSU
7	32	80.0	298	V317_HSV1
8	31	77.5	106	Y002_MEDEA
9	31	77.5	215	Y473_AQURE
10	31	77.5	218	F5PF_MOUSE
11	31	77.5	273	O5SR_CANFA
12	31	77.5	291	BHCL_RHO2
13	31	77.5	292	P1RPG_PSEU1
14	31	77.5	299	BPHC_PSEPA
15	31	77.5	393	HENK_ECOLI
16	31	77.5	484	LGGA_BACSU
17	31	77.5	747	MPEA_MOUSE
18	31	77.5	748	MPEA_RAT
19	31	77.5	922	NRP1_RAT
20	31	77.5	923	NRP1_HUMAN
21	31	77.5	923	NRP1_MOUSE
22	31	1045	1045	PRT5_SERMA
23	31	77.5	1045	PRT5_SERMA
24	30	75.0	236	RIP3_SAROF
25	30	75.0	253	RIP5_SAROF
26	30	75.0	253	RIP7_SAROF
27	30	75.0	292	RIP2_SAROF
28	30	75.0	293	RIP2_DIACA
29	30	75.0	299	RIP6_SAROF
30	30	75.0	357	5H5A_HUMAN
31	30	75.0	357	5H5A_MOUSE
32	30	75.0	357	5H5A RAT
33	30	75.0	370	5H5B_MOUSE

## ALIGNMENTS

RESULT 1  
HXKL\_BOVIN  
ID HXKL\_BOVIN  
STANDARD; PRT; 918 AA.  
AC P2159;  
DT 01-AUG-1992 (Rel. 23, Created)  
DT 01-AUG-1992 (Rel. 23, Last sequence update)  
DT 20-AUG-2001 (Rel. 40, Last annotation update)  
DE HEXOKINASE, TYPE I (EC 2.7.1.1) (HK I) (BRAIN FORM HEXOKINASE).  
GN HK1.  
OS Bos taurus (Bovine).  
OC Eukaryota; Metacozia; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae; Bovinae; Bos.  
OC OX NCBI\_TaxID:9913;  
RN [1]  
RP SBQUENCE FROM N.A.  
RX MEDLINE#89247146; Pubmed=2719857;  
RA Griffin L.D., Macgregor G.R., Muzny D.M., Harter J., Cook R.G., McCabe E.R.;  
RT "Synthesis and characterization of a bovine hexokinase 1 cDNA probe by mixed oligonucleotide primed amplification of cDNA using high complexity primer mixtures";  
RL Biochem. Med. Metab. Biol. 41:125-131(1989).  
RP [2]  
RX SEQUENCE FROM N.A.  
RA Griffin L.D., Gaib B.D., Wheeler D.A., Davison D., Adams V., McCabe E.R.;  
RT "Mammalian hexokinase 1: evolutionary conservation and structure to function analysis";  
RL Genomics 11:1014-1024(1991).  
CC -I CATALYTIC ACTIVITY: HEXOKINASE IS AN ALLOSTERIC ENZYME INHIBITED BY ITS PRODUCT GLC-6-P.  
CC ..- PATHWAY: FIRST STEP OF SEVERAL METABOLIC PATHWAYS.  
CC ..- SUBUNIT: MONOMER.  
CC -I SURCELLULAR LOCATION: BOUND TO THE OUTER MITOCHONDRIAL MEMBRANE.  
CC TTS HYDROPHOBIC N-TERMINAL SEQUENCE MAY BE INVOLVED IN MEMBRANE BINDING.  
CC ..- MISCELLANEOUS: IN VERTEBRATES THERE ARE FOUR MAJOR GLUCOSE-PHOSPHORYLATING ISOENZYMES, DESIGNATED HEXOKINASE I, II, III AND IV (GLUCOKINASE).  
CC ..- SIMILARITY: THE N- AND C-TERMINAL HALVES OF THIS HEXOKINASE SHOW EXTENSIVE SEQUENCE SIMILARITY TO EACH OTHER. THE CATALYTIC ACTIVITY IS ASSOCIATED WITH THE C-TERMINUS WHILE REGULATORY FUNCTION IS ASSOCIATED WITH THE N-TERMINUS.  
CC ..- SIMILARITY: BELONGS TO THE HEXOKINASE FAMILY.

This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

DR	EMBL: M6510; AAA51661.1; -.
DR	HSSP: P19367; IIPKB.
DR	InterPro: IPR01312; Hexokinase.
DR	Pfam: PF00349; hexokinase; 2.
DR	PRINTS: PRO0475; HEXOKINASE.
DR	PRODOM: PD001109; Hexokinase; 2.
DR	PROSITE: PS00378; HEXOKINASES; 2.
KW	Transferase; Kinase; Glycolysis; Allosteric enzyme; Repeat;
KW	ATP-transferase; Membrane; ATP-binding; Membrane.
FT	DOMAIN 1 12 HYDROPHOBIC.
FT	DOMAIN 13 475 REGULATORY (BY SIMILARITY).
FT	DOMAIN 476 918 CATALYTIC (BY SIMILARITY).
FT	DOMAIN 149 175 GLUCOSE-BINDING (POTENTIAL).
FT	DOMAIN 597 623 GLUCOSE-BINDING (POTENTIAL).
FT	NP_BIND 84 89 ATP (POTENTIAL).
FT	NP_BIND 537 558 ATP (POTENTIAL).
FT	BINDING 558 558 ATP (POTENTIAL).
SQ	SEQUENCE 918 AA; 103064 MW; 1DCBFB7ID06FEEB6 CRC64;

n.

RESULT	2
Y052_HUMAN	
ID	Y052_HUMAN
STANDARD;	PRT;
P42285;	1045 AA.
DT	01-NOV-1995 (Rel. 32, Created)
DT	20-AUG-2001 (Rel. 40, Last annotation update)
DE	HYPOTHETICAL PROTEIN KIAA0052 (FRAGMENT).
GN	KIAA0052.
OS	Homo sapiens (Human).
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX	NCBI_TAXID=9606;
RN	[1]
RP	SSEQUENCE FROM N.A.
RC	TISSUE-Bone marrow;
RX	MEDLINE=96051398; PubMed=7584044;
RA	Nomura N., Nagase T., Miyajima N., Sazuka T., Tanaka A., Saeki N., Kawarabayashi Y., Ishikawa K.-I., Tabata S.;
RT	'Prediction of the coding sequences of unidentified human genes. The coding sequences of 40 new genes (KIAA041-KIAA080) determined by analysis of cDNA clones from human cell line KG-1.'; DNA Res. 1:223-229 (1994).
RL	- - - SUBCELLULAR LOCALIZATION: NUCLEAR (POTENTIAL).
CC	- - - SIMILARITY: BELONGS TO THE SKI2 SUBFAMILY OF HELICASES.
CC	- - - CAUTION: THIS IS A CONCEPTUAL TRANSLATION: A FRAMESHIFT INTRODUCED IN POSITION 837 TO MAKE THE PROTEIN LONGER AND THE SIMILARITY WITH SKI2.
CC	-----
CC	This SWISS-PROT entry is copyright. It is produced through the European Bioinformatics Institute. There are no restrictions on non-profit institutions as long as its content is modified and this statement is not removed. Usage by and entities requires a license agreement (see <a href="http://www.isb-sib.ch">http://www.isb-sib.ch</a> or send an email to <a href="mailto:license@isb-sib.ch">license@isb-sib.ch</a> ).
CC	-----
DR	EMLB; D29541; BAM06124 1; ALT_FRAME.
DR	InterPro; IPR0110; DEAD.
DR	InterPro; IPR01650; Helicase_C.
DR	Pfam; PF00270; DEAD; 1.
DR	Pfam; PF00271; helicase_C; 1.
SMART; SM00487; DEXDC; 1.	SMART; SM00490; HELIC_C; 1.
CC	-----

[1] —  
 RQ SEQUENCE FROM N.A.  
 RP STRAIN=5288C / F11679;  
 RX MEDLINE#96287653; PUBMED#6686380;  
 RA Nasr F.; Bécam A.-M.; Herbert C.J.;  
 RT "the sequence of 36.8 kb from the left arm of chromosome XIV reveals  
 RT 24 complete open reading frames; 18 correspond to new genes, one of  
 RT which encodes a protein similar to the human myotonic dystrophy  
 RT kinase";  
 RL Yeast 12:169-175(1996).  
 RN [2]  
 RP SQUENCE OF 1-65 FROM N.A.  
 RC STRAIN=S288C;  
 RX MEDLINE#96109932; PubMed#8619318;  
 RA Mallet L.; Bustereau F.; Jacquot M.;  
 RT "A 43.5 kb segment of yeast chromosome XIV, which contains MPA2,  
 RT MP2, CAP/SRV2, NAM9, FKB1/FPR1/RBP1, MOM22 and CPR1, predicts an  
 RT adenosine deaminase gene and 14 new open reading frames";  
 RL Yeast 11:1195-1209(1995).  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>).  
 CC -----  
 DR EML; X92517; CA063290; 1; .  
 DR EML; 271425; CA096037; 1; .  
 DR EML; 271424; CA096033; 1; .  
 DR SGD; S0005033; YNL149C.  
 KW Hypothetical protein.  
 SQ SEQUENCE 129 AA; 15053 MW; ALA41E3DABCEAA06 CRC64;  
 Query Match 85.0%; Score 34; DB 1; Length 129;  
 Best Local Similarity 100.0%; Pred. No. 19;  
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 2 FGWKGK 6  
 DB 84 FGWKGK 88

**RESULT 4**

GP52\_HUMAN STANDARD; PRT; 361 AA.

ID GPS2\_HUMAN STANDARD; PRT; 361 AA.

AC 09Y2TS; OT5654; PRT; 361 AA.

DT 20-AUG-2001 (Rel. 40, Created)

DT 20-AUG-2001 (Rel. 40, Last sequence update)

DE PROBABLE G PROTEIN-COUPLED RECEPTOR GPR52.

GN GPR52.

OS Homo sapiens (Human)

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

OX NCBI\_TaxID=9606;

RN [1]

RP SEQUENCE FROM N.A.

RX MEDLINE=99132305; PUBMED=9931487;

RA Sawczarzko M., Nguyen T., Lee D.K., Lynch K.R., Cheng R., Heng H.H.Q., George S.R., O'Dowd B.F.; "Identification and cloning of three novel human G protein-coupled receptor genes GPR52, PSIGPR53 and GPR55: GPR55 is extensively expressed in human brain.";

RT Brain Res. Mol. Brain Res. 64:193-198(1999).

RN [2]

RP SEQUENCE FROM N.A.

RA Wilkinson J.; Submitted (JUN-1998) to the EMBL/GenBank/DDJB databases.

CC -I- FUNCTION: ORPHAN RECEPTOR.

CC -I- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN.

CC -I- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.

This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

CC DR InterPro; IPR000276; GPCR\_Rhodopsin.

DR PFAM; PF00001; 7tm\_1; 1.

DR PRINTS; PR00237; GPERRHODOPSN.

DR PROSITE; PS00217; G\_PROTEIN\_RECEP\_FL1; 1.

DR PROSITE; PS50262; G\_PROTEIN\_RECEP\_FL2; 1.

KW G-protein coupled receptor; Transmembrane; Glycoprotein.

FT DOMAIN 1 44 EXTRACELLULAR (POTENTIAL).

FT DOMAIN 45 65 1 (POTENTIAL).

FT DOMAIN 66 87 CYTOPLASMIC (POTENTIAL).

FT TRANSMEM 88 108 2 (POTENTIAL).

FT DOMAIN 109 116 EXTRACELLULAR (POTENTIAL).

FT TRANSMEM 117 137 3 (POTENTIAL).

FT DOMAIN 138 159 CYTOPLASMIC (POTENTIAL).

FT TRANSMEM 160 180 4 (POTENTIAL).

FT DOMAIN 181 203 EXTRACELLULAR (POTENTIAL).

FT TRANSMEM 204 224 5 (POTENTIAL).

FT DOMAIN 225 265 CYTOPLASMIC (POTENTIAL).

FT TRANSMEM 266 286 6 (POTENTIAL).

FT DOMAIN 287 317 EXTRACELLULAR (POTENTIAL).

FT TRANSMEM 318 361 CYTOPLASMIC (POTENTIAL).

FT CARBOHYD 2 2 N-LINKED (GLCNAC. . .) (POTENTIAL).

FT CARBOHYD 13 13 N-LINKED (GLCNAC. . .) (POTENTIAL).

FT CARBOHYD 20 20 N-LINKED (GLCNAC. . .) (POTENTIAL).

FT CONFLICT 21 21 A -> V (IN REF. 2).

FT CONFLICT 64 64 A -> V (IN REF. 2).

FT CONFLICT 115 115 R -> O (IN REF. 2).

FT CONFLICT 307 307 V -> I (IN REF. 2).

FT CONFLICT 322 322 G -> S (IN REF. 2).

FT CONFLICT 332 332 F -> S (IN REF. 2).

SQ SEQUENCE 361 AA: C935B47C8F27B5A CRC54;

**RESULT 5**

TKTC\_CRAPL STANDARD; PRT; 519 AA.

ID TKTC\_CRAPL

AC 042675; PRT; 519 AA.

DT 01-NOV-1997 (Rel. 35, Created)

DT 01-NOV-1997 (Rel. 35, Last sequence update)

DE TRANSKETOLASE, CHLOROPLAST (EC 2.2.1.1) (TK) (FRAGMENT).

RT TKT3.

RA Craterostigma plantagineum.

CC OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Asteridae; euasterids I; Lamiales; Lamiales incertae sedis; Toreniae; OC Craterostigma.

CX NCBI\_TaxID=4153;

RN [1]

RP SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.

RX MEDLINE=95163594; PUBMED=859749;

RA Barnaccia G., Schwall G., Lottspeich F., Salamini F., Bartels D.; RT "The transketolase gene family of the resurrection plant: rehydration phase." ENBO J. 14:510-518 (1995).

CC -I- CATALYTIC ACTIVITY: SEDOHEPTULOSE 7-PHOSPHATE + D-GLYCERALDEHYDE 3-PHOSPHATE = D-RIBOSE 5-PHOSPHATE + D-XYLULOSE 5-PHOSPHATE.

CC -I- CODACTOR: THIAMINE PYROPHOSPHATE.

CC -I- SUBUNIT: HOMODIMER.

CC -I- TISSUE SPECIFICITY: CONSTITUTIVELY EXPRESSED IN LEAVES AND ROOTS.

CC -I- SIMILARITY: BELONGS TO THE TRANSKETOLASE FAMILY.

This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

CC DR EMBL; Z46646; CAA86607; 1.

DR HSSP; P23254; ITRKA.

DR InterPro; IPR003160; Transketolase.

DR PFAM; PF00456; transketolase; 1.

DR PROSITE; PS00801; TRANSKETOLASE\_1; PARTIAL.

DR PROSITE; PS00802; TRANSKETOLASE\_2; 1.

KW Transferase; Thiamine pyrophosphate; Chloroplast; Multigene family.

FT NON\_TER 1 1

SQ SEQUENCE 519 AA: 56187 MW: 7AFDD34CE2EC56B5 CRC64;

Query Match Best local Similarity Score 85.0%; 34; DB 1; Length 519; Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 FGWGK 6

Db 473 FGWGK 477

**RESULT 6**

TKT\_BACSU STANDARD; PRT; 667 AA.

ID TKT\_BACSU





This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

CC EMBL; AF031533; AAB86633.1; -.

CC GCRDB; GCR\_2500; -.

CC InterPro; IPR00276; GPCR\_Rhodopsin.

CC InterPro; IPR001760; Opsin.

CC Pfam; PF00001; 7tm\_1; 1.

CC PROSITE; PS00237; G-PROTEIN\_RECEP\_FL1\_1; 1.

CC PROSITE; PS00238; OPSIN; PARTIAL.

KW Photoreceptor; Retinal protein; Transmembrane; Glycoprotein; Vision; Phosphorylation; Lipoprotein; Palmitate; G-protein coupled receptor.

FT NON\_TER 1 1

FT DOMAIN <1 5

FT TRANSMEM 6 30

FT DOMAIN 31 42

FT TRANSMEM 43 68

FT DOMAIN 69 82

FT TRANSMEM 83 102

FT DOMAIN 103 121

FT TRANSMEM 122 145

FT DOMAIN 146 171

FT TRANSMEM 172 199

FT DOMAIN 200 221

FT TRANSMEM 222 245

FT DOMAIN 246 253

FT TRANSMEM 254 273

FT DISULFID 79 156

FT BINDING 265 265

FT NON\_TER 273 273

SQ SEQUENCE 30373 MW: C1A27C95D1649F85 CRC64;

Query Match 77.5%; Score 31; DB 1; Length 273; Best Local Similarity 80.0%; Pred. No. 1.2e+02; Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 2 FGWGR 6

Db 142 FGWGR 146

RESULT 12

BHCL\_RHOGO STANDARD; PRT; 291 AA.

AC P47731; 47731

DT 01-FEB-1996 (Rel. 33, Created) 15-JUL-1999 (Rel. 38, Last annotation update)

DE BIPHENYL-2,3-DIOL 1,2-DIOXYGENASE I (EC 1.13.11.39) (230HBP OXYGENASE 1) (2,3-DIHYDROXYBIPHENYL DIOXYGENASE I) (DHBD I).

GN Rhodococcus globerulus.

OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae; Actinomycetales; Corynebacterineae; Nocardiaceae; Rhodococcus.

OC OX NCBI\_TaxID=3008; [1]

RN SEQUENCE FROM N.A.

RP MEDLINE=94171820; PubMed=8126007;

RX STRAIN=P6; MEDLINE=94171820; PubMed=8126007;

RA Asturias J.A., Eltis L.D., Prucha M., Timmels K.N.; Rhodococcus globorubris sp. Identification of a new family of extradiol dioxygenases.; J. Biol. Chem. 269:7807-7815(1994);

RA "Analysis of three 2,3-dihydroxybiphenyl 1,2-dioxygenases found in Rhodococcus globorubris sp. Identification of a new family of extradiol dioxygenases." J. Biol. Chem. 269:7807-7815(1994);

RA "CATALYTIC ACTIVITY: BIPHENYL-2,3-DIOL + O(2) = 2-HYDROXY-6-OXO-6-PHENYLHEXA-2,4-DIENOATE + H(2)O.

CC CC

CC COFACTOR: FERROUS ION.

CC PATHWAY: DEGRADATION OF BIPHENYLS AND POLYCHLOROBIPHENYLS (PCB) TO BENZOIC ACID AND CHLOROBENZOIC ACIDS.

CC SIMILARITY: BELONGS TO THE EXTRADIOL RING-CLEAVAGE DIOXYGENASE FAMILY.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

CC EMBL; X75633; CA53397.1; -.

CC HSSP; P47228; IHNAN.

CC InterPro; IPR000486; Extradiol\_dioxygenase.

CC PROSITE; PS00082; EXTRADIOL\_DIOXYGENASE\_1.

CC Pfam; PF000903; Glyoxalase; 1.

CC PRODOM; PD00097; Extradiol\_dioxygenase; 1.

CC KW Oxidoreductase; Dioxygenase; Aromatic hydrocarbons catabolism; Iron.

FT METAL 146 146

FT METAL 210 210

FT METAL 260 260

FT SEQUENCE 291 AA; 32081 MW; 104F189FEI6DDA CRC64;

Query Match 77.5%; Score 31; DB 1; Length 291; Best Local Similarity 80.0%; Pred. No. 1.2e+02; Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 DFGWG 5

Db 260 EFGWG 264

RESULT 13

BPHC\_PSE1 STANDARD; PRT; 292 AA.

AC P17297; 052441; 1

DT 01-AUG-1990 (Rel. 15, Created) 01-NOV-1995 (Rel. 32, Last annotation update)

DT 15-JUL-1999 (Rel. 38, Last annotation update)

DE BIPHENYL-2,3-DIOL 1,2-DIOXYGENASE (EC 1.13.11.39) (230HBP OXYGENASE) (2,3-DIHYDROXYBIPHENYL DIOXYGENASE) (DHBD).

DE BPHC\_PSE1

GN Pseudomonas sp. (strain KKS102).

OC Bacteria; Proteobacteria.

OX NCBI\_TaxID=307;

RN [1]

RP SEQUENCE FROM N.A.

RX MEDLINE=89213365; PubMed=2540155;

RA Kimbara K., Hashimoto T., Fukuda M., Koana T., Takagi M., Oishi M., Yano K.;

RA "Cloning and sequencing of two tandem genes involved in degradation of 2,3-dihydroxybiphenyl to benzoic acid in the polychlorinated biphenyl-degrading soil bacterium Pseudomonas sp. strain KKS102." J. Bacteriol. 171:2740-2747(1989).

RA [2]

RP X-RAY CRYSTALLOGRAPHY (1.8 ANGSTROMS).

RX MEDLINE=96226036; PubMed=8636975;

RA Senda T., Sugiyama K., Narita H., Yamamoto T., Kimbara K., Fukuda M., Sato M., Yano K., Mitsui Y.;

RA "Three-dimensional structures of free form and two substrate complexes of an extradiol ring-cleavage type dioxygenase, the BphC enzyme from Pseudomonas sp. strain KKS102." J. Mol. Biol. 255:735-752(1996).

RP "CATALYTIC ACTIVITY: BIPHENYL-2,3-DIOL + O(2) = 2-HYDROXY-6-OXO-6-PHENYLHEXA-2,4-DIENOATE + H(2)O.

CC CC

CC COFACTOR: FERROUS ION

CC PATHWAY: DEGRADATION OF BIPHENYLS AND POLYCHLOROBIPHENYLS (PCB) TO BENZOIC ACID AND CHLOROBENZOIC ACIDS.

CC -!- SUBUNIT: HOMOCTAMER.  
 CC -!- SIMILARITY: BELONGS TO THE EXTRADIOL RING-CLEAVAGE DIOXYGENASE FAMILY.

CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

CC -----  
 DR EMBL: M26433; AAA25150.1; --.  
 DR EMBL: D17319; BAA04141.1; --.  
 DR PIR: A32312; DAPSIC.  
 DR PDB: 1DHY; 15-OCT-95.  
 DR IPR00486; Extradiol\_dioxygenase.  
 DR InterPro; IPR00325; Glyoxalase\_1.  
 DR Pfam; PF00903; Glyoxalase; 2.  
 DR PROSITE; PS00082; EXTRADIOL\_DIOXYGENAS; 1.  
 DR KW Oxidoreductase; Dioxxygenase; Aromatic hydrocarbons catabolism; Iron; 3D-structure.  
 FT INIT\_MET 0  
 FT METAL 145 145 IRON.  
 FT METAL 209 209 IRON.  
 FT METAL 260 260 IRON.  
 SQ SEQUENCE 292 AA; 32113 MW; 5D7D912F79EA8476 CRC64;

Query Match 77.5%; Score 31; DB 1; Length 292;  
 Best local Similarity 80.0%; Pred. No. 1 2e+02; Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Oy 1 DRGWG 5  
 Db 260 EFGWG 264

RESULT 14

BPHC\_PSEPA ID BPHC\_PSEPA STANDARD: PRT; 299 AA.  
 AC P11122; DT 01-JUL-1989 (Rel. 11, Created)  
 DT 01-JUL-1989 (Rel. 11, Last sequence update)  
 DT 15-JUL-1999 (Rel. 38, Last annotation update)  
 DE BIPHENYL-2,3-DIOL 1,2-DIOXYGENASE (EC 1.13.11.39) (230HBP OXYGENASE) (2,3-DIHYDROXYBIPHENYL DIOXYGENASE) (DHBD).  
 GN BPHC.  
 OS Pseudomonas paucimobilis (Sphingomonas paucimobilis).  
 OC Bacteria; Proteobacteria; alpha subdivision; Sphingomonadaceae;  
 OC Sphingomonas.  
 OC NCBI\_TaxID=13689;  
 RN [1] SEQUENCE FROM N.A.  
 RP STRAIN="Q1;  
 RC MEDLINE=8826911; PubMed=3137968;  
 RA Taira K., Hayase N., Arimura N., Yanashita S., Miyazaki T.,  
 RA Furukawa K.;  
 RT "Cloning and nucleotide sequence of the 2,3-dihydroxybiphenyl dioxygenase gene from the PCB-degrading strain of Pseudomonas paucimobilis Q1";  
 RT Biochemistry 27:3990-3996(1998).  
 CC -!- CATALYTIC ACTIVITY: BIPHENYL-2,3-DIOL + O(2) = 2-HYDROXY-6-OXO-BENZOIC ACID AND CHLOROBENZOIC ACIDS.  
 CC -!- COFACTOR: FERROUS ION  
 CC -!- PATHWAY: DEGRADATION OF BIPHENYLS AND POLYCHLOROBIPHENYLS (PCB) TO  
 CC BENZOIC ACID AND CHLOROBENZOIC ACIDS.  
 CC -!- SUBUNIT: HOMOCTAMER.  
 CC -!- SIMILARITY: BELONGS TO THE EXTRADIOL RING-CLEAVAGE DIOXYGENASE FAMILY.

This SWISS-PROT entry is copyright. It is produced through a collaboration

CC -----  
 CC -!- SUBUNIT: HOMOCTAMER.  
 CC -!- SIMILARITY: BELONGS TO THE EXTRADIOL RING-CLEAVAGE DIOXYGENASE FAMILY.

CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

CC -----  
 DR EMBL: M20640; AAA25678.1; --.  
 DR EMBL: A28718; A28718.  
 DR HSSP; P1729; 10M.  
 DR InterPro; IPR00486; Extradiol\_dioxygenase.  
 DR InterPro; IPR00325; Glyoxalase\_1.  
 DR PFM; PR00903; Glyoxalase; 2.  
 DR PRODOM; PD00977; Extradiol\_dioxygenase; 1.  
 DR PROSITE; PS00082; EXTRADIOL\_DIOXYGENAS; 1.  
 DR KW Oxidoreductase; Dioxxygenase; Aromatic hydrocarbons catabolism; Iron.  
 FT METAL 149 149 IRON (BY SIMILARITY).  
 FT METAL 212 212 IRON (BY SIMILARITY).  
 FT METAL 263 263 IRON (BY SIMILARITY).  
 SQ SEQUENCE 299 AA; 33095 MW; 29746CAB79D6FC7B CRC64;

Query Match 77.5%; Score 31; DB 1; Length 299;  
 Best local Similarity 80.0%; Pred. No. 1 3e+02; Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Oy 1 DRGWG 5  
 Db 263 EFGWG 267

RESULT 15

HEMX\_ECOLI ID HEMX\_ECOLI STANDARD: PRT; 393 AA.  
 AC F09127; DT 01-MAR-1989 (Rel. 10, Created)  
 DT 01-MAR-1989 (Rel. 10, Last sequence update)  
 DT 20-AUG-2001 (Rel. 40, Last annotation update)  
 DE POTATIVE UROPORPHIRIN III C-METHYLTRANSFERASE (EC 2.1.1.107) (UROGEN TIT METHYLASE) (ORF X).  
 GN HEMX OR B3803.  
 OS Escherichia coli.  
 OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;  
 OC Escherichia  
 OX NEBI\_TaxID=562;  
 RN [1] SEQUENCE FROM N.A.  
 RC STRAIN=K12;  
 RA MEDLINE=89098348; PubMed=3062586;  
 RA Sasarman A., Echelard Y., Letowski J., Tardif D., Drolet M.;  
 RT "Nucleotide sequence of the hemX gene, the third member of the Uro operon of Escherichia coli K12.";  
 RL Nucleic Acids Res. 16:11835-11835(1988).  
 RN [2] SEQUENCE FROM N.A.  
 RP STRAIN=K12 / CS520;  
 RC MEDLINE=89041586; PubMed=3054815;  
 RA Alefounder P.R.;  
 RT "The sequence of hemC, hemD and two additional E. coli genes.;"  
 RT Nucleic Acids Res. 16: 9871-9871(1988).  
 RN [3] SEQUENCE FROM N.A.  
 RC STRAIN=K12 / MG1655;  
 RX MEDLINE=92338234; PubMed=1279743;  
 RA Daniels D.L., Plunkett G., III, Burland V.D., Blattner F.R.;  
 RT "Analysis of the Escherichia coli genome: DNA sequence of the region from 84.5 to 86.5 minutes;"  
 RL Science 257:771-778(1992).  
 RN [4] SEQUENCE OF 1-12.  
 RP STRAIN=K12 / EMG2;  
 RC MEDLINE=9743975; PubMed=9298646;  
 RX Link A.J., Robison K., Church G.M.;

RT "Comparing the predicted and observed properties of proteins encoded  
in the genome of *Escherichia coli* K-12.";  
RL Electrophoresis 18(1259-1313)(1997).  
CC -1- CATALYTIC ACTIVITY: 2-S-ADENOSYL-L-METHIONINE + UROPORPHYRIN III  
= 2-S-ADENOSYL-L-HOMOCYSTEINE + SIROHYDROCHLORIN.  
CC -!- PATHWAY: INVOLVED IN THE BIOSYNTHESIS OF SIROHEM AND COBALAMIN.  
CC  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
between the Swiss Institute of Bioinformatics and the EMBL outstation  
the European Bioinformatics Institute. There are no restrictions on its  
use by non-profit institutions as long as its content is in no way  
modified and this statement is not removed. Usage by and for commercial  
entities requires a license agreement. (See <http://www.isb-sib.ch/announce/>  
or send an email to license@isb-sib.ch).  
CC  
CC  
DR EMBL; X13406; CAA31172.1; -  
DR EMBL; X12614; CAA31134.1; -  
DR EMBL; M87049; AAA67599.1; -  
DR EMBL; AE000456; AAC16806.1; -  
DR PIR; S02185;  
DR PIR; S30693; S30693;  
DR ECODBASE; B043.0; 6TH EDITION.  
DR ECODBASE; B043.1; 6TH EDITION.  
DR Ecogene; EG10433; hemX.  
KW Porphyrin biosynthesis; Transferase; Methyltransferase;  
KW Complete proteome;  
SEQUENCE 393 AA;  
SQ 42963 MW; 9D272C6401D0E354 CRC64;

Query Match 77.5%; Score 31; DB 1; Length 393;  
Best Local Similarity 80.0%; Pred. No. 1.6e+02; Mismatches 0;  
Matches 4; Conservative 1; Indels 0; Gaps 0;  
QY 2 FGWGK 6  
Db :|||  
55 YGWGK 59

Search completed: November 5, 2001, 12:58:55  
Job time: 116 sec

GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 5, 2001, 12:56:59 ; Search time 22.88 Seconds  
(without alignments)  
38.358 Million cell updates/sec

Title: US-08-894-356C-21  
Perfect score: 40 DFGWGX 6

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 473505 seqs, 146272329 residues

Total number of hits satisfying chosen parameters: 473505

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : SPREMBL\_17:\*

1: sp\_archea:\*

2: sp\_bacteria:\*

3: sp\_fungi:\*

4: sp\_human:\*

5: sp\_invertebrate:\*

6: sp\_mammal:\*

7: sp\_minc:\*

8: sp\_organelle:\*

9: sp\_phage:\*

10: sp\_plant:\*

11: sp\_rabbit:\*

12: sp\_virus:\*

13: sp\_vertebrate:\*

14: sp\_unclassified:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

**SUMMARIES**

Result No.	Score	Query Match Length	DB ID	Description
1	40	100.0	219	10 Q9M5K9
2	40	100.0	434	10 Q9FH97
3	40	100.0	435	10 Q9FYMO
4	40	100.0	436	10 Q9FYMI
5	40	100.0	439	10 Q9ZTK5
6	40	100.0	442	10 Q9FF43
7	40	100.0	443	10 Q9SNQ9
8	40	100.0	443	10 Q9FJNO
9	40	100.0	446	10 Q9MBC1
10	40	100.0	448	10 Q9MBD4
11	40	100.0	448	10 Q9FTD1
12	40	100.0	448	10 Q9FH98
13	40	100.0	449	10 Q9LJB4
14	40	100.0	450	10 Q9SNEO
15	40	100.0	451	10 Q94470
16	40	100.0	455	10 P93094
17	40	100.0	455	10 Q9LU88
18	40	100.0	458	10 Q9LJA2
19	40	100.0	460	10 Q43583

**ALIGNMENTS**

RESULT	1
Q9M5K9	PRELIMINARY; PRT; 219 AA.
ID Q9M5K9;	
AC Q9M5K9;	
DT 01-OCT-2000 (TREMBLrel. 15, Created)	
DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)	
DT 01-OCT-2000 (TREMBLrel. 15, Last annotation update)	
DE P21J9, 20-LIKE PROTEIN (FRAGMENT).	
OS Euphorbia esula (Leafy spurge).	
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicots; Rosidae; eurosids I; Malpighiales; Euphorbiaceae; Euphorbia.	
OX NCBI_TaxID:3993;	
RN [1]	
RP SEQUENCE FROM N.A.	
RC TISSUE-UNDERGROUND ADVENTITIOUS BUDS;	
RA Anderson J.V., Horvath D.P.;	
RT "Identification of mRNAs expressed in underground adventitious buds of Euphorbia esula (leafy spurge).";	
RT Submitted (JAN-2000) to the EMBL/GenBank/DDBJ databases.	
DR EMBL; AF227981; AAF34801.1; -.	
PT NON-TER 1	
SO SEQUENCE . 219 AA; 24236 MW; DDD96c93FC87BC30 CRC64;	
Query Match Similarity 100.0%; Score 40; DB 10; Length 219; Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;	
QY 1 DFGWGX 6	
DB 156 DFGWGX 161	
RESULT 2	
DFN97 ID Q9FH97 PRELIMINARY; PRT; 434 AA.	
AC Q9FH97;	
DT 01-MAR-2001 (TREMBLrel. 16, Created)	
DT 01-MAR-2001 (TREMBLrel. 16, Last sequence update)	
DT 01-JUN-2001 (TREMBLrel. 17, Last annotation update)	
DE ANTHRANILATE N-HYDROXYCINNAMOYL/BENZOYLTRANSFERASE-LIKE PROTEIN.	

OS Arabidopsis thaliana (Mouse-ear cress).  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;  
 OC eurosids II; Brassicales; Brassicaceae; Arabidopsis.  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=COLUMBIA;  
 RX MEDLINE=20181125; PubMed=10718197;  
 RA Sato S., Nakamura Y., Kaneko T., Katoh T., Asamizu E., Kotani H.,  
 RA Tabata S.;  
 RT "Structural analysis of Arabidopsis thaliana chromosome 5. X. Sequence  
 features of the regions of 3,076,755 bp covered by sixty P1 and TAC  
 clones.";  
 RT DNA Res. ?; 31:63(2000);  
 RL EMBL: AB020742; BAB10950.1; -.  
 DR InterPro: IPR003480; Transferase.  
 DR Pfam: PF02458; Transferase; 1.  
 KW Transferase.  
 SQ SEQUENCE 434 AA; 48635 MW; 7E4C4C61DA1624AA CRC64;

Query Match 100.0%; Score 40; DB 10; Length 434;  
 Best Local Similarity 100.0%; Pred. No. 18; Mismatches 0;  
 Matches 6; Conservative 0; Indels 0; Gaps 0;  
 Qy 1 DFGWCK 6  
 Db 376 DFGWCK 381

RESULT 3  
 Q9FMM0 PRELIMINARY; PRT; 435 AA.  
 ID Q9FMM0;  
 DT 01-MAR-2001 (TREMBLrel. 16, created)  
 DT 01-MAR-2001 (TREMBLrel. 16, last sequence update)  
 DE F21J9.9  
 OS Arabidopsis thaliana (Mouse-ear cress).  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;  
 OC eurosids II; Brassicales; Brassicaceae; Arabidopsis.  
 OC NCBI-TaxID=3702;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Johnson-Hopson C., Brooks S., Buehler E., Chao Q., Khan S., Kim C.,  
 RA Shinn P., Altafai H., Bel O., Chin C., Chiou J., Choi E., Conn L.,  
 RA Conway A., Hansen N., Howing B., Koo T., Lam B., Lee J.,  
 RA Lenz C., Li J., Liu A., Liu S., Mukharsky N., Nguyen M.,  
 RA Palm C., Pham P., Sakano H., Schwartz J., Southwick A., Thaveri A.,  
 RA Moriumi M., Vaysberg M., Yu G., Rederspiel N.A., Theologis A.,  
 RA Ecker J.R.;  
 RT "Genomic sequence for Arabidopsis thaliana BAC F21J9 from chromosome  
 RT I.";  
 RL Submitted (DEC-1996) to the EMBL/GenBank/DBJ databases.  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RA Ecker J.R.;  
 RL Submitted (FEB-1997) to the EMBL/GenBank/DBJ databases.  
 RP SEQUENCE FROM N.A.  
 RA Ecker J.R.;  
 RL Submitted (MAR-1997) to the EMBL/GenBank/DBJ databases.  
 RN [4]  
 RP SEQUENCE FROM N.A.  
 RA Ecker J.R.;  
 RL Submitted (MAR-1997) to the EMBL/GenBank/DBJ databases.  
 RN [5]  
 RP SEQUENCE FROM N.A.  
 RA Ecker J.R.;  
 RL Submitted (JUN-1997) to the EMBL/GenBank/DBJ databases.  
 RN [6]

Query Match 100.0%; Score 40; DB 10; Length 435;  
 Best Local Similarity 100.0%; Pred. No. 18; Mismatches 0;  
 Matches 6; Conservative 0; Indels 0; Gaps 0;  
 Qy 1 DFGWCK 6  
 Db 369 DFGWCK 374

RESULT 4  
 Q9FYML PRELIMINARY; PRT; 436 AA.  
 ID Q9FYML;  
 AC Q9FYML;  
 DT 01-MAR-2001 (TREMBLrel. 16, created)  
 DT 01-MAR-2001 (TREMBLrel. 16, last sequence update)  
 DE F21J9.8  
 OS Arabidopsis thaliana (Mouse-ear cress).  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;  
 OC eurosids II; Brassicales; Brassicaceae; Arabidopsis.  
 OC NCBI-TaxID=3702;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Johnson-Hopson C., Brooks S., Buehler E., Chao Q., Khan S., Kim C.,  
 RA Shinn P., Altafai H., Bel O., Chin C., Chiou J., Choi E., Conn L.,  
 RA Conway A., Hansen N., Howing B., Koo T., Lam B., Lee J.,  
 RA Lenz C., Li J., Liu A., Liu K., Liu S., Mukharsky N., Nguyen M.,  
 RA Palm C., Pham P., Sakano H., Schwartz J., Southwick A., Thaveri A.,  
 RA Moriumi M., Vaysberg M., Yu G., Rederspiel N.A., Theologis A.,  
 RA Ecker J.R.;  
 RT "Genomic sequence for Arabidopsis thaliana BAC F21J9 from chromosome  
 RT I.";  
 RL Submitted (DEC-1996) to the EMBL/GenBank/DBJ databases.  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RA Ecker J.R.;  
 RL Submitted (FEB-1997) to the EMBL/GenBank/DBJ databases.  
 RP SEQUENCE FROM N.A.  
 RA Ecker J.R.;  
 RL Submitted (MAR-1997) to the EMBL/GenBank/DBJ databases.  
 RN [4]  
 RP SEQUENCE FROM N.A.  
 RA Ecker J.R.;  
 RL Submitted (MAR-1997) to the EMBL/GenBank/DBJ databases.  
 RN [5]  
 RP SEQUENCE FROM N.A.  
 RA Ecker J.R.;  
 RL Submitted (JUN-1997) to the EMBL/GenBank/DBJ databases.

RP SEQUENCE FROM N.A.  
 RA Ecker J.R.;  
 RL Submitted (JUL-2000) to the EMBL/GenBank/DBJ databases.  
 RN [8]  
 RP SEQUENCE FROM N.A.  
 RA Cheuk R., Shinn P., Brooks S., Buehler E., Chao Q., Johnson-Hopson C.,  
 RA Khan S., Kim C., Altafai H., Bel O., Chin C., Chiou J., Choi E.,  
 RA Conn L., Conway A., Hansen N., Howing B., Koo T., Lam B.,  
 RA Lee J., Lenz C., Li J., Liu A., Liu S., Mukharsky N.,  
 RA Nguyen M., Palm C., Pham P., Sakano H., Schwartz J., Southwick A.,  
 RA Thaveri A., Toriumi M., Vaysberg M., Yu G., Davis R., Federspiel N.,  
 RA Theologis A., Ecker J.;  
 DR Submitted (AUG-2000) to the EMBL/GenBank/DBJ databases.  
 DR EMBL: AC000103; AACF97979.1; -.  
 DR InterPro: IPR003480; Transferase.  
 DR Pfam: PF02458; Transferase; 1.  
 SQ SEQUENCE 435 AA; 49288 MW; B369CEEC11D982DF CRC64;

RN [6]  
 RP SEQUENCE FROM N A.  
 RA Ecker J.R.;  
 RL Submitted (JUL-2000) to the EMBL/GenBank/DDBJ databases.  
 RN [7]  
 RP SEQUENCE FROM N A.  
 RA Ecker J.R.;  
 RL Submitted (JUL-2000) to the EMBL/GenBank/DDBJ databases.  
 RN [8]  
 RP Cheuk R., Shinn P., Brooks S., Buehler E., Chao Q., Johnson-Hopson C., Khan S., Kim C., Alzafai H., Bai B., Chin C., Chiu J., Choi E., Conn L., Conway A., Gonzalez A., Hansen N., Howing B., Koo T., Lam B., Lee J., Lenz C., Li J., Liu A., Liu J., Liu S., Mukharyk N., Nguyen M., Palm C., Pham P., Sakano H., Schwartz J., Southwick A., Thaveri A., Toriumi M., Vaysberg M., Yu G., Davis R., Federspiel N., RA Theologis A., Ecker J.;  
 RL Submitted (AUG-2000) to the EMBL/GenBank/DDBJ databases.  
 DR EMBL; AC000103; AF091955.1; -;  
 DR InterPro; IPR003480; Transferase.  
 DR Pfam; PF02458; Transferase; 1.  
 SQ SEQUENCE 436 AA; 48824 MW; B303A52DE3FD1355 CRC64;

Query Match 100.0%; Score 40; DB 10; Length 436;  
 Best Local Similarity 100.0%; Pred. No. 18;  
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DFGWPK 6  
 Db 372 DFGWPK 377

RESULT 5  
 Q9ZTK5 PRELIMINARY; PRT; 439 AA.

AC Q9ZTK5;  
 DT 01-MAY-1999 (TREMBREL. 10, Last sequence update)  
 DT 01-JUN-2001 (TREMBREL. 17, Last annotation update)  
 DE DEACYLIVINDOLINE 4-O-ACETYLTRANSFERASE (EC 2.3.1.107).  
 DAT.  
 GN CATHARANTHUS ROSEUS (Rosy periwinkle) (Madagascar periwinkle).  
 OS Eukaryota; Viridiplantae; Streptophytina; Embryophytina; Tracheophytina; Spermatophytina; Magnoliophytina; eudicotyledons; core eudicots; Asteridae; euasterids I; Gentianales; Apocynaceae; Catharanthus.  
 OC NCBI\_TAXID=4058;  
 OX [1]  
 RN SEQUENCE FROM N A.  
 RC STRAIN=CV; LITTLE DELICATA;  
 RX MEDLINE=98346012; PUBMED=9681034;  
 RA St-Pierre B., Laflamme P., Alarco A.M., De Luca V.;  
 RT "The terminal O-acetyltransferase involved in vindoline biosynthesis defines a new class of proteins responsible for coenzyme A-dependent acyl transfer";  
 RT Plant J. 14:703-713(1998).  
 RL EMBL; AF053307; AAC93311.1; -;  
 DR InterPro; IPR003480; Transferase.  
 DR Pfam; PF02458; Transferase; 1.  
 KW Transferase; Acyltransferase.  
 SQ SEQUENCE 439 AA; 49873 MW; 1357C76550E0D04D CRC64;

Query Match 100.0%; Score 40; DB 10; Length 439;  
 Best Local Similarity 100.0%; Pred. No. 18;  
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DFGWPK 6  
 Db 380 DFGWPK 385

RESULT 6  
 Q9FF43 PRELIMINARY; PRT; 442 AA.  
 AC Q9FF43;  
 DT 01-MAR-2001 (TREMBREL. 16, Created)  
 DT 01-JUN-2001 (TREMBREL. 17, Last sequence update)  
 DR AMTHRANILATE N-HYDROXYCINNAMOYL/BENZOYLTRANSFERASE-LIKE PROTEIN.  
 OS Arabidopsis thaliana (Mouse-ear cress);  
 RA Spermatophytina; Viridiplantae; Streptophytina; Embryophytina; Tracheophytina; eudicotyledons; core eudicots; Rosidae; OC eurosids II; Brassicales; Brassicaceae; Arabidopsis.  
 OX NCBI\_TAXID=3702;  
 RN [1]  
 RP SEQUENCE FROM N A.  
 RA STRAIN=COLUMBIA;  
 RX MEDLINE=97471969; PubMed=9330910;  
 RA Mato S., Kotani H., Nakamura Y., Kaneko T., Asamizu E., Fukami M., Miyajima N., Tabata S.;  
 RT "Structural analysis of Arabidopsis thaliana chromosome 5. I. Sequence features of the 1.6 Mb regions covered by twenty physically assigned P1 clones";  
 RL DNA RES. 4:215-230(1997).  
 DR EMBL; AB00247; BAB11280.1; -;  
 DR InterPro; IPR003480; Transferase.  
 DR Pfam; PF02458; Transferase; 1.  
 KW SEQUENCE 442 AA; 49491 MW; 1D280C885875DA47 CRC64;

Query Match 100.0%; Score 40; DB 10; Length 442;  
 Best Local Similarity 100.0%; Pred. No. 18;  
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DFGWPK 6  
 Db 384 DFGWPK 389

RESULT 7  
 Q9 SND9 PRELIMINARY; PRT; 443 AA.

AC Q9 SND9;  
 DT 01-MAY-2000 (TREMBREL. 13, Created)  
 DT 01-MAY-2000 (TREMBREL. 13, Last sequence update)  
 DT 01-JUN-2001 (TREMBREL. 17, Last annotation update)  
 DE ANTHRANILATE N-HYDROXYCINNAMOYL/BENZOYLTRANSFERASE-LIKE PROTEIN.  
 GN F11C1120.  
 OS Arabidopsis thaliana (Mouse-ear cress).  
 OC Eukaryota; Viridiplantae; Streptophytina; Embryophytina; Tracheophytina; Spermatophytina; Magnoliophytina; eudicotyledons; core eudicots; Rosidae; OC eurosids II; Brassicales; Brassicaceae; Arabidopsis.  
 OX NCBI\_TAXID=3702;  
 RN [1]  
 RP SEQUENCE FROM N A.  
 RA Barrios M., Colmado M.C., Navarro P., Terol J., Perez-Alonso M., Mewes H.W., Lemke K., Mayer K.F.X., Quettier F., Salanoubat M.;  
 RL Submitted (NOV-1999) to the EMBL/GenBank/DDBJ databases.  
 RN [2]  
 RP SEQUENCE FROM N A.  
 RA EU Arabidopsis sequencing project;  
 RL Submitted (DEC-1999) to the EMBL/GenBank/DDBJ databases.  
 DR EMBL; AL132976; CAB6207.1; -;  
 DR InterPro; IPR003480; Transferase.  
 DR Pfam; PF02458; Transferase; 1.  
 KW SEQUENCE 443 AA; 49960 MW; C4ADB67A0E3AB3D4 CRC64;

Query Match 100.0%; Score 40; DB 10; Length 443;  
 Best Local Similarity 100.0%; Pred. No. 18;  
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DFGWPK 6

RESULT	8	Best-Local Similarity 100.0%; Pred- No. 19; Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0					
Q9FUND		PRELIMINARY;	PRT;	443 AA.			
ID	Q9FJNO;						
AC	Q9FJNO_;	01-MAR-2001 (TREMBLrel. 16, Created)					
DT	01-MAR-2001 (TREMBLrel. 16, Last sequence update)						
DE	N-HYDROXYCINNAMAMOYLTRANSFERASE.						
OS	Arabidopsis thaliana (Mouse-ear cress);						
OC	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;						
OC	Spermato phyta; Magnoliophyta; eudicots; core eudicots; Rosidae;						
OC	eurosid s II; Brassicales; Brassicaceae; Arabidopsis.						
OX	NCBI_TAXID=3702;						
RN	[1]						
RP	SEQUENCE FROM N.A.						
RX	STRAIN=COLUMBIA; MEDLINE=98403884; PubMed=9734815;						
RA	Kotani H., Nakamura Y., Saito S., Asanizu E., Kaneko T., Miyajima N., Tabata S.;						
RT	"Structural analysis of Arabidopsis thaliana chromosome 5. VI. Sequence features of the regions of 1,367,185 bp covered by 19 physically assigned P1 and TAC clones.";						
RT	DNA Res. 5;203:215(1998).						
RL	EMBL: AB013395; BAA08854.1; -						
DR	InterPro: IPR003480; Transferase.						
DR	Pfam: PF02458; Transferase; 1.						
KW	Transferase.						
SQ	SEQUENCE 443 AA; 49950 MW; 16EFC01B99035AAD CRC64;						
Query Match Best Local Similarity 100.0%; Score 40; DB 10; Length 443; Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0							
Q9MBC1		PRELIMINARY;	PRT;	446 AA.			
ID	Q9MBC1						
AC	Q9MBC1_;	01-OCT-2000 (TREMBLrel. 15, Created)					
DT	01-OCT-2000 (TREMBLrel. 15, Last sequence update)						
DE	ANTHOCYANIN ACYLTRANSFERASE (FRAGMENT).						
OS	Perilla frutescens.						
OC	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;						
OC	Magnoliophyta; eudicots; core eudicots;						
OC	Asteridae; eudicotids I; Lamiales; Lamiaceae; Perilla.						
OX	NCBI_TAXID=48386;						
RN	[1]						
RP	SEQUENCE FROM N.A.						
RC	TISSUE=LEAF;						
RA	Sakakibara K.Y., Tanaka Y., Mizutani M.F., Fujiwara H., Fukui Y., Ashikari T., Yamaguchi M., Kusumi T.;						
RT	"Molecular and biochemical characterization of a novel hydroxycinnamamoyl-CoA: anthocyanin 3-O-glucoside-6-"; Submitted (JUN-1999) to the EMBL/GenBank/DBJ databases.						
RL	EMBL: AB029340; BAA93475.1; -						
DR	InterPro: IPR003480; Transferase.						
DR	Pfam: PF02458; Transferase; 1.						
KW	Transferase; Acyltransferase.						
FT	NON-TER 1						
SQ	SEQUENCE 446 AA; 50675 MW; DC83C69AF248A593 CRC64;						
Query Match 100.0%; Score 40; DB 10; Length 446; Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0							
Q9MBD4		PRELIMINARY;	PRT;	448 AA.			
ID	Q9MBD4						
AC	Q9MBD4_;	01-OCT-2000 (TREMBLrel. 15, Created)					
DT	01-OCT-2000 (TREMBLrel. 15, Last sequence update)						
DE	ANTHOCYANIN ACYLTRANSFERASE-LIKE PROTEIN.						
OS	Arabidopsis thaliana (Mouse-ear cress);						
OC	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;						
OC	Spermato phyta; Magnoliophyta; eudicots; core eudicots; Rosidae;						
OC	eurosid s III; Brassicales; Brassicaceae; Arabidopsis.						
OX	NCBI_TAXID=3702;						
RN	[1]						
RP	SEQUENCE FROM N.A.						
RC	STRAIN=COLUMBIA; MEDLINE=915623; PubMed=10049488;						
RA	Asanizu E., Saito S., Kaneko T., Nakamura Y., Kotani H., Miyajima N., Tabata S.;						
RT	"Structural analysis of Arabidopsis thaliana chromosome 5. VIII. Sequence features of the regions of 1,081,958 bp covered by seventeen physically assigned P1 and TAC clones.";						
RL	DNA Res. 5;379:391(1998).						
DR	EMBL: AB016892; BAB10831.1; -						
DR	InterPro: IPR003480; Transferase.						
DR	Pfam: PF02458; Transferase; 1.						

KW Transferase; Acyltransferase.  
 SQ *SEQUENCE* 448 AA; 49916 MW; 27878BBEBB8984E CRC64;

Query Match 100.0%; Score 40; DB 10; Length 448;  
 Best Local Similarity 100.0%; Pred. No. 19;  
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 DFGWCK 6  
 Db 393 DFGWCK 398

RESULT 12  
 Q9FH98 PRELIMINARY; PRT; 448 AA.  
 ID Q9FH98  
 AC Q9FH98;  
 DT 01-MAR-2001 (TREMBLrel. 16, Last sequence update)  
 DT 01-MAY-2001 (TREMBLrel. 16, Last sequence update)  
 DE ANTHRANILATE N-HYDROXYCINNAMOYL/BENZOYLTRANSFERASE-LIKE PROBIN.  
 OS Arabidopsis thaliana (Mouse-ear cress).  
 OC Eukaryota; Viridiplantae; Streptophytina; Embryophytina; Tracheophytina; Spermatophytina; Magnoliophytina; eudicotyledons; core eudicots; Rosidae; eurosids II; Brassicales; Brassicaceae; Arabidopsis.  
 OC NCBI\_TaxID=3702;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=COLUMBIA;  
 RX MEDLINE-20181125; PubMed=10718197;  
 RA Sato S., Nakamura Y., Kaneko T., Katoh T., Asamizu E., Kotani H., Tabata S.;  
 RT "Structural analysis of Arabidopsis thaliana chromosome 5. X. Sequence features of the regions of 3,076,755 bp covered by sixty PI and TAC clones";  
 RT DNA Res. 7:31-63 (2000).  
 RL EMBL; AB020742; BAB10949\_1; -.  
 DR InterPro; IPR003480; Transferase.  
 DR Pfam; PF02458; Transferase; 1.  
 KW SEQUENCE 448 AA; 49939 MW; 04EB676545093DBA CRC64;

Query Match 100.0%; Score 40; DB 10; Length 449;  
 Best Local Similarity 100.0%; Pred. No. 19;  
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 DFGWCK 6  
 Db 389 DFGWCK 394

RESULT 13  
 Q9LJB4 PRELIMINARY; PRT; 449 AA.  
 ID Q9LJB4  
 AC Q9LJB4;  
 DT 01-OCT-2000 (TREMBLrel. 15, Created)  
 DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)  
 DT 01-JUN-2001 (TREMBLrel. 17, Last annotation update)  
 DE ANTHOCYANIN 5-AROMATIC ACYLTRANSFERASE-LIKE PROTEIN.  
 OS Arabidopsis thaliana (Mouse-ear cress).  
 OC Eukaryota; Viridiplantae; Streptophytina; Embryophytina; Tracheophytina; Spermatophytina; Magnoliophytina; eudicotyledons; core eudicots; Rosidae; eurosids II; Brassicales; Brassicaceae; Arabidopsis.  
 OC NCBI\_TaxID=3702;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA EU Arabidopsis sequencing project;  
 RL Submitted (DEC-1999) to the EMBL/GenBank/DDBJ databases.  
 DR EMBL; AL132976; CAB62306\_1; -.  
 DR InterPro; IPR003480; Transferase.  
 DR Pfam; PF02458; Transferase; 1.  
 KW Transferase.  
 SQ *SEQUENCE* 450 AA; 49805 MW; 5D4A4AACB3F0C6 CRC64;

Query Match 100.0%; Score 40; DB 10; Length 450;  
 Best Local Similarity 100.0%; Pred. No. 19;  
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 DFGWCK 6  
 Db 391 DFGWCK 396

RESULT 15  
 Q9LJB4 PRELIMINARY; PRT; 451 AA.  
 ID Q9LJB4  
 AC Q9LJB4;  
 DT 01-AUG-1998 (TREMBLrel. 07, Created)  
 DT 01-AUG-1998 (TREMBLrel. 07, Last sequence update)  
 DT 01-AUG-2001 (TREMBLrel. 07, Last sequence update)  
 DE PUTATIVE ANTHRANILATE N-HYDROXYCINNAMOYL/BENZOYLTRANSFERASE.  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=COLUMBIA;  
 RA Kaneko T., Kato T., Sato S., Nakamura Y., Asamizu E., Tabata S.; Submitted (OCT-1999) to the EMBL/GenBank/DDBJ databases.  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=COLUMBIA;

RX Pubmed=10907853;  
 RA Nakamura Y.;  
 RT Structural analysis of Arabidopsis thaliana chromosome 3. II. Sequence features of the regions of 4,251,695 bp covered by ninety PI,  
 RT TAC and BAC clones.",  
 RT DNA Res. 7:217-221(2000).  
 DR EMBL; AP000606; BAB01191\_1; -.  
 DR InterPro; IPR003480; Transferase.  
 DR Pfam; PF02458; Transferase; 1.  
 KW Transferase; Acyltransferase.  
 SQ *SEQUENCE* 449 AA; 49439 MW; 2446D502AEC9D99B CRC64;

RX Pubmed=10907853;  
 RA Nakamura Y.;  
 RT Structural analysis of Arabidopsis thaliana chromosome 3. II. Sequence features of the regions of 4,251,695 bp covered by ninety PI,  
 RT TAC and BAC clones.",  
 RT DNA Res. 7:217-221(2000).  
 DR EMBL; AP000606; BAB01191\_1; -.  
 DR InterPro; IPR003480; Transferase.  
 DR Pfam; PF02458; Transferase; 1.  
 KW Transferase; Acyltransferase.  
 SQ *SEQUENCE* 449 AA; 49439 MW; 2446D502AEC9D99B CRC64;

OC eurosids II; Brassicales; Brassicaceae; Arabidopsis.  
OX NCBI\_TaxID=3702;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=CV. COLUMBIA;  
RA Rounseley S.D., Kaul S., Lin X., Ketchum K.A., Crosby M.L.,  
RA Brandon R.C., Sykes S.M., Mason T.M., Kerlavage A.R., Adams M.D.,  
RA Somerville C.R., Venner J.C.;  
RL Submitted (JUN-1998) to the EMBL/GenBank/DBJ databases.  
DR EMBL: AC002392; AAC120251; -.  
DR InterPro: IPR03480; Transferase.  
DR Pfam: PF02458; Transferase; 1.  
KW Transferase; 451 AA; 50434 MW; B13B2CBF4D4B4482 CRC64;  
SQ

Query Match 100.0%; Score 40; DB 10; Length 451;  
Best Local Similarity 100.0%; Pred. No. 19;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 DFGWCK 6  
||| |||  
Db 397 DFGWCK 402

Search completed: November 5, 2001, 12:58:39  
Job time: 100 sec